UNIVERSITY OF CALIFORNIA
SANTA CRUZ

FLEXIBLE DYNAMIC QUANTILE LINEAR MODELS

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

STATISTICS AND APPLIED MATHEMATICS

by

Raquel A. Barata

December 2021

The Dissertation of Raquel A. Barata
is approved:

Raquel Prado, Co-chair

Bruno Sansó, Co-chair

Athanasios Kottas

Gabriel Huerta

Peter F. Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures **vi**

List of Tables **xiv**

Abstract **xviii**

Acknowledgments **xxi**

1 Introduction **1**

2 Fast Inference for Time-Varying Quantiles via Flexible Dynamic Models **8**
2.1 Introduction and Background 9
2.2 The exDQLM 12
2.2.1 Markov chain Monte Carlo algorithm 13
2.2.2 Importance sampling variational Bayes algorithm 15
2.2.3 Comparison criteria 19
2.2.4 Discount factor selection 20
2.2.5 Comments on prior selection and inference of σ and γ . . . 21
2.3 Simulation study 22
2.3.1 Results 25
2.4 Conclusion 29

3 Estimating a Non-linear Response in a Time-Varying Quantile with Application to the Characterization of Atmospheric Rivers **32**
3.1 Introduction 32
3.2 Transfer Function exDQLM 36
3.2.1 MCMC and ISVB Algorithm Augmentations 37
3.3 Estimating the 0.85 quantile IVT threshold 38
3.3.1 Model selection 38
3.3.2 IVT analysis 42
3.3.3 Comparison with the DQLM special case 48
3.3.4 Model Validation 52

3.4 Conclusion . . . 55

4 exdqlm: An R Package for Estimation and Analysis of Flexible Dynamic Quantile Linear Models 57
4.1 Introduction . . . 57
4.2 Extended dynamic quantile linear models . . . 58
4.2.1 Prior specification . . . 59
4.2.2 Posterior estimation . . . 59
4.2.3 Specification of the evolution covariance via discount factors 61
4.2.4 k-step-ahead forecast . . . 62
4.2.5 Transfer function model . . . 62
4.2.6 Special cases . . . 64
4.3 Model Diagnostics . . . 65
4.4 Examples . . . 66
4.4.1 Lake Huron . . . 67
4.4.2 Sunspots . . . 73
4.4.3 Big Tree water flow . . . 79
4.5 Conclusion . . . 88

5 Joint Estimation of Time-Varying Quantiles via Flexible Multivariate Dynamic Models 89
5.1 Background . . . 89
5.1.1 The multivariate exAL . . . 91
5.1.2 Specifying $\mathbf{s}$ and $\mathbf{v}$ . . . 92
5.2 The MVexDQLM . . . 93
5.2.1 MCMC algorithm . . . 96
5.2.2 ISVB algorithm . . . 96
5.3 Comparison criteria . . . 98
5.4 Big Tree water flow analysis . . . 101
5.4.1 Modeling framework . . . 102
5.4.2 Results . . . 105
5.5 Multivariate IVT analysis . . . 110
5.5.1 Modeling Framework . . . 110
5.5.2 Incorporating spatial correlation . . . 115
5.5.3 Results . . . 116
5.6 Conclusion . . . 124

6 Conclusion 125

A Algorithm Details 128
A.1 MCMC Forward Filtering Backwards Sampling . . . 128
A.2 ISVB Forward Filtering Backwards Smoothing . . . 130
A.3 exDQLM ISVB Importance Sampling . . . 132

A.4 exDQLM ISVB Closed Form Integrals 133
A.5 MVexDQLM ISVB Importance Sampling 134
A.6 MVexDQLM ISVB Closed Form Integrals 138
A.7 Multivariate IVT analysis additional results 139

List of Figures

2.1 Density functions of $\text{exAL}_{p_0}(y|\gamma, \mu, \sigma)$ defined in Equation (2.3) with $\mu = 0$, $\sigma = 1$ and different values of γ for fixed quantiles $p_0 = 0.05$, 0.5, and 0.85. The black solid line corresponds to the AL density, which is a special case when $\gamma = 0$. 11

2.2 Simulation study datasets. From left to right are the data simulated from the Stochastic Volatility (SV) model, exDQLM, and generalized DLM (gDLM) described in Section 2.3. 22

2.3 MCMC results. Smoothed posterior distributions of dynamic quantiles $p_0 = 0.05, 0.5, 0.85$. exDQLM estimates in blue, DQLM in red, and DLM in orange. Dotted lines indicate the 95% CrI from the smoothed posterior distributions and solid lines indicate posterior mean estimates. Due to the scale of the gDLM data, we focus on time periods which highlight the similarities and/or differences of the models at the three different quantiles. 26

2.4 ISVB and MCMC comparison results. Posterior distributions of dynamic quantiles $p_0 = 0.05, 0.5, 0.85$. MCMC exDQLM estimates in blue, ISVB exDQLM estimates in purple. Dotted lines indicate the 95% CrI from the smoothed posterior distributions and solid lines indicate posterior mean estimates. Due to the scale of the gDLM data, we focus on period of only 100 time points and omit the median estimates for visual clarity. 27

3.1 Top panel: Average daily IVT magnitude in the grid cell containing Santa Cruz, CA. The dashed vertical lines enclose the two time periods enlarged in the middle panel. Middle panel: A closer look at the two time periods highlighted in the top panel. Years 1982 to 1985 illustrate a time period which saw an exceptional amount of rain. Years 2012 to 2015 illustrate a period which was exceptionally dry. ARs detected by the scheme proposed in Guan and Waliser (2015) in the Santa Cruz grid cell are indicated with dark triangles. ARs detected in neighboring coastal grid cells are indicated with lighter circles. Bottom panel: ELI anomalies resulting from the de-seasonalization of the interpolated ELI. The dashed, horizontal line is at zero, for reference. 33
3.2 M_{01} seasonal components for harmonics $l = 1, 2, \ldots, 6$. 41
3.3 M_1 retained seasonal components for harmonics $l = 1, 2, 4$. 41
3.4 Top panel: IVT data with overlaid solid line indicating the M_1 MAP 0.85 quantile threshold of the average daily IVT magnitude in Santa Cruz, CA. The vertical dashed lines enclose the two periods enlarged in the bottom panels. 42
3.5 Top panel: Effects of ELI captured by the transfer function component, ζ_t. Bottom panel: Instantaneous effects of ELI, ψ_t. In both panels, dark grey lines indicate the MAP estimates. 95% CrI are indicated by the grey shaded regions. Dashed horizontal dashed lines are at zero, for reference. Left vertical dashed lines enclose years 1982 to 1985 in which CA saw an exceptional amount of rain. Right vertical dashed lines enclose 2012 to 2015 in which CA was exceptionally dry. 43

3.6 60-step-ahead quantile forecast beginning November 2, 2019 through December 31, 2019 overlaid on the IVT magnitude data. The vertical dot-dashed line is at the beginning of the forecast period, November 2, 2019, for reference. From M_1: Solid lines indicate mean estimates and dotted lines 95% CrI, with the filtered estimates seen leading up to November 2$^{\text{nd}}$ and the forecast estimates beyond November 2$^{\text{nd}}$. From M_0: The dashed lines indicate the mean estimates, again with the filtered and forecast estimates seen before and after November 2$^{\text{nd}}$, respectively. 45

3.7 Top panel: In purple, M_1 MAP 0.85 quantile threshold of the average daily IVT magnitude in Santa Cruz, CA. In red, M_2 MAP 0.85 quantile. The blue vertical lines enclose a period in which CA saw an exceptional amount of rain. The red vertical lines enclose a period in which CA was exceptionally dry. Bottom panels: Years 1982 to 1985 (left, exceptional amount of rain) and years 2012 to 2015 (right, exceptionally dry). The MAP 0.85 quantile from M_1 and M_2 are seen in purple and red, respectively, with 95% CrIs indicated with dashed lines. 50

3.8 Top panel: Effects of ELI captured by the transfer function component, ζ_t. Bottom panel: Instantaneous effects of ELI, ψ_t. In both panels, M_1 is seen in purple and M_2 is seen in red. Solid lines indicate the MAP estimates. 95% CrI are indicated by the shaded regions. Dashed horizontal orange lines are at zero, for reference. Blue vertical lines enclose years 1982 to 1985 in which CA saw an exceptional amount of rain. Red vertical lines enclose 2012 to 2015 in which CA was exceptionally dry. 51

3.9 60-step-ahead quantile forecast beginning November 2, 2019 through December 31, 2019. From M_1: The solid magenta line indicates the forecast means and the dashed magenta indicate the 95% CrI of the forecasted 0.85 quantile. Also included are the filtered means and 95% CrI up to November 2, 2019 in purple solid and dashed lines, respectively. From M_2: The solid orange line indicates the forecast means and the dashed orange lines the 95% CrIs. The filtered means and 95% CrI up to November 2, 2019 in red solid and dashed lines, respectively. The IVT magnitude data is seen in grey. . . . 52
3.10 QQ-plots of the transformed one-step-ahead predictive distribution functions. M_0 is seen in green, M_1 in purple, and M_2 in red. . . . 53
3.11 ACF plots of the one-step-ahead predictive distribution functions. M_0 is seen in green, M_1 in purple, and M_2 in red. 53
3.12 MAP estimates of the standard one-step-ahead predictive forecast errors. M_0 is seen in green, M_1 in purple, and M_2 in red. 55

4.1 Example 1: Lake Huron. Trace plot (left) and density (right) of the MCMC samples of the skewness parameter γ estimated at the median. 70
4.2 Example 1: Lake Huron. Left: MAP estimates and 95% CrIs of the estimated quantiles, plotted with the data (grey). Right: Forecasted quantile estimates and 95% credible intervals (seen after 1972), along with the filtered quantile estimates and 95% credible intervals (seen from 1952 to 1972). 72
4.3 Example 2: Sunspots. Left: The sunspot time series from 1700 to 1988. Center: MAP estimates and 95% CrIs of the estimated quantiles from the DQLM (red) and exDQLM (blue), plotted with the data (grey) from 1750 to 1850. Right: Histogram of samples from the approximated posterior distribution of γ. 76

4.4 Example 2: Sunspots. The QQ-plots (left column), ACF plots (center column) and standard forecast errors (right column) from the DQLM (red) and exDQLM (blue). 78
4.5 Example 3: Big Tree water flow. Top: Average monthly water flow at the Big Tree gauge of the San Lorenzo river in Santa Cruz County, CA, plotted on the log-scale. Bottom: Monthly values of Niño 3.4. Both time series span January 1937 through April 2021. 80
4.6 Example 3: Big Tree water flow. Top: MAP estimates and 95% CrIs of the estimated quantiles from the models with the regression component (purple) and transfer function component (green), plotted with the data (grey) from 1970 to 1990. Middle: MAP estimates and 95% CrIs of the annual seasonal components. Bottom: MAP estimates and 95% CrIs of the components which model the quantile association with Niño 3.4. The estimates of the dynamic regression component $X_t\boldsymbol{\theta}_{4,t}$ of M_1 where $\boldsymbol{\theta}_{4,t}$ denotes the fourth element of the state vector, are seen in purple. The estimates of the transfer function component ζ_t of M_2 are in green. The horizontal orange dotted line is at zero for reference. 83
4.7 Example 3: Big Tree water flow. MAP estimates and 95% CrIs of parameters ζ_t and ψ_t from the transfer function component. Horizontal dotted orange lines are at zero for reference. 85
4.8 Example 3: Big Tree water flow. 18-step-ahead quantile forecast beginning October 2019 through April 2021 overlaid on the log Big Tree water flow time series. The vertical dot-dashed line is at the beginning of the forecast period, October 2019, for reference. Solid lines indicate mean estimates and dotted lines 95% CrI, with the filtered estimates seen leading up to October 2019 and the forecast estimates beyond October 2019. Estimates from M_1, the quantile regression model, are seen in purple/pink. Estimates from M_2, the transfer function model, are seen in green. 87

5.1 Corr(y_i, y_j) for case (c); common s and iid v_j, for varying values of γ_i (x-axes), γ_j (y-axes), p_0 (columns) and R_{ij} (rows). 94
5.2 Corr(y_i, y_j) for case (d); iid s_j and iid v_j, for varying values of γ_i (x-axes), γ_j (y-axes), p_0 (columns) and R_{ij} (rows). 94
5.3 Top panel: Monthly water flow on the log-scale (cfs) at the Big Tree water gauge on the San Lorenzo river in Santa Cruz county, CA. Middle panel: Monthly precipitation (in) recorded in Santa Cruz county. Bottom panel: Monthly temperature (F) also recorded in Santa Cruz county. All data span from January 1937 to February 2021. 102
5.4 Data with overlaid solid line indicating the M_1 MAP quantiles with surrounding dashed lines indicated in the 95% CrIs. Top panel: log-Big Tree water flow data with 0.15-quantile estimates. Middle panel: Precipitation data with 0.15-quantile estimates. Bottom panel: Temperature data with 0.85-quantile estimates. 107
5.5 Baseline (first column) and seasonal (second column) components of the quantiles estimated by M_0 (purple) and M_0 (green). Solid lines indicate the MAP estimates with surrounding dashed lines indicated in the 95% CrIs. Top row: log-Big Tree water flow first order-polynomial baseline and seasonal component including harmonic $l = 1$. Middle row: Precipitation first order-polynomial baseline and seasonal component including harmonics $l = 1, 2$. Bottom row: Temperature first order-polynomial baseline and seasonal component including harmonics $l = 1, 2, 3$. 108

5.6 12-step-ahead quantile forecasts for the Big Tree log-water flow, precipitation, and temperature data (left to right). Forecasts begin February 2020 through February 2021 and are overlaid on the respective time-series. The vertical dot-dashed line is at the beginning of the forecast period, February 2020, for reference. Solid lines indicate mean estimates and dotted lines 95% CrI, with the filtered estimates seen leading up to February 2020 and the forecast estimates beyond February 2020. Results from M_1 are seen in purple and magenta, and results from M_0 are seen in dark and light green. 109

5.7 IVT magnitude observation locations spanning all of CA. Triangles indicate coastal locations at which Guan and Waliser (2015) detect ARs to make landfall. The dark green filled triangle marks the location nearest Santa Cruz CA. The light green filled triangles mark the locations nearest Fort Bragg (in Northern CA) and Santa Barbara (in Southern CA). Finally, the locations outlined in blue and red are those we label as climatologically Northern and Southern CA, respectively. 111

5.8 Average daily IVT magnitude near Fort Bragg CA (FB; top panel), Santa Cruz CA (SC; middle panel), and Santa Barbara CA (SB; bottom panel), spanning from 1979 through 2015. The dashed vertical lines enclose the years 1982 to 1985, a time period in which CA saw an exceptional amount of rain, and years 2012 to 2015, a period which was exceptionally dry. 112

5.9 Fixed values of $\boldsymbol{\sigma}$ displayed by latitude (y-axis) and longitude (x-axis). 114

5.10 MAP estimates of $\boldsymbol{\gamma}$ displayed by latitude (y-axis) and longitude (x-axis). 117

5.11 IVT data with overlaid solid line indicating the MAP 0.85-quantile estimates near Fort Bragg CA (FB; top panel), Santa Cruz CA (SC; middle panel), and Santa Barbara CA (SB; bottom panel). The dashed vertical lines enclose the years 1982 to 1985 and 2012 to 2015, time periods in which CA was exceptionally wet and dry, respectively. 118

5.12 MAP estimates of the second-order polynomial trend components displayed by latitude (y-axes) and longitude (x-axes). The starting value at $t = 1$, $\eta_{1,1}$ (top), and change over the entire time period, $13505\eta_2$ (bottom). 119

5.13 MAP estimates of the annual (top), semi-annual (middle), and quarterly (bottom) amplitudes displayed by latitude (y-axes) and longitude (x-axes). 121

5.14 Top panel: Common effects of ELI captured by the transfer function component, ζ_t. Bottom panel: Common instantaneous effects of ELI, ψ_t. In both panels, dark grey lines indicate the MAP estimates. 95% CrI are indicated by the grey shaded regions. Dashed horizontal dashed lines are at zero, for reference. Left vertical dashed lines enclose years 1982 to 1985 in which CA saw an exceptional amount of rain. Right vertical dashed lines enclose 2012 to 2015 in which CA was exceptionally dry. 122

5.15 First two PCs of the components of the the forecast covariances which correspond to Northern CA (left) and Southern CA (right). Green dots indicate the PCs for times at which the considered region of CA is experiencing "high AR activity". 123

A.1 MAP estimates of the annual (top), semi-annual (middle), and quarterly (bottom) phases displayed by latitude (y-axes) and longitude (x-axes). 140

List of Tables

2.1 Posterior summaries for γ and σ (where applicable): mean (95% CrI). Mean check loss of the MAP dynamic quantile. KL divergences of the one-step-ahead distributions. Posterior predictive loss criterion (pplc) under the check loss function. Computation runtime (min). 30

2.2 Median proportion of coverage agreement between the 95% CrIs from the ISVB and MCMC algorithms. That is, the proportion of the 95% CrIs from the MCMC algorithm covered by the 95% CrIs from the ISVB algorithm. 31

3.1 IVT analysis results for M_{01} and M_1. Optimal δ_ζ, δ_ψ, and $\tilde{\lambda}$. Posterior summaries (format: mean (95% CrI)) for skewness parameter γ, baseline value at $t = 1$, $\eta_{1,1}$, as well as the change in the baseline at each time step, η_2. pplc: Posterior predictive loss criterion under the check loss function. KL: Kullback-Liebler divergences of the one-step-ahead distributions. 40

3.2 IVT analysis results for M_0 and M_1. Optimal δ_ζ, δ_ψ, and $\tilde{\lambda}$, as discussed in Section 3.3.1. Posterior summaries (format: mean (95% CrI)) for skewness parameter γ, baseline at $t = 1$ $\eta_{1,1}$, the change in the baseline at each time step η_2, annual amplitude A_1, annual phase P_1, semi-annual amplitude A_2, semi-annual phase P_2, quarterly amplitude A_4, quarterly phase P_4. pplc: Posterior predictive loss criterion under the check loss function. KL: Kullback-Liebler divergences of the one-step-ahead distributions. Run-time: ISVB run-times in minutes. 44

3.3 IVT analysis results for M_1 and M_2. Identical values of (δ_ζ, δ_ψ) used in both models and optimized within M_1. Rate parameter $\tilde{\lambda}$ optimized within each model. The scale parameter σ is fixed for M_1 as discussed in Section 2.7, and the skewness parameter γ is fixed to 0 for M_2 (reducing the model to the DQLM). Posterior summaries (format: mean (95% CrI)) for M_2 scale parameter σ, M_1 skewness parameter γ, baseline at $t = 1$ $\eta_{1,1}$, the change in the baseline at each time step η_2, annual amplitude A_1, annual phase P_1, semi-annual amplitude A_2, semi-annual phase P_2, quarterly amplitude A_4, quarterly phase P_4. pplc: Posterior predictive loss criterion under the check loss function. KL: Kullback-Liebler divergences of the one-step-ahead distributions. Run-time: ISVB run-times in minutes. 49

4.1 Translation from mathematical symbols to parameters used **exdqlm** functions, as well as default values where applicable. 59

4.2 Translation from mathematical symbols to parameters used in functions `exdqlmMCMC` and `exdqlmISVB`, as well as default values. Here $\mathbf{I}_2$ denotes the identity matrix of dimension 2. 61

4.3 Translation from mathematical symbols to additional parameters used in function `exdqlmForecast`. 62

4.4 Translation from mathematical symbols to additional parameters used in function `transfn_exdqlmISVB`. 64
4.5 Example 3: Big Tree water flow. Model diagnostic output from `exdqlmChecks` and resulting values. 88

5.1 Corr(y_i, y_j) for the four possible modeling options of latent parameters v_j and s_j. 92
5.2 Big Tree log-water flow, precipitation, and temperature analysis results for M_0 and M_1. Posterior summaries (format: mean (95% CrI)) for skewness parameters γ_{BT} (BT water flow), γ_{prec} (precipitation), γ_{temp} (temperature), as well as the latent correlations $\mathbf{R}_{\text{BT,prec}}$ (BT flow and precipitation), $\mathbf{R}_{\text{BT,temp}}$ (BT flow and temperature), $\mathbf{R}_{\text{prec,temp}}$ (precipitation and temperature). pplc: Posterior predictiveloss criterion under the check loss function. KL: Kullback-Liebler divergences of the marginal one-step-ahead distributions. Run-time: MCMC run-times in hours. 106
5.3 Multivariate IVT analysis results. Posterior summaries for exponential correlation function parameter ϕ in the format: mean (95% CrI). Optimal values of $\tilde{\lambda}, \delta_\zeta, \delta_\psi$. Run-time: ISVB run-time in hours. 115
5.4 95% CrI estimates of $\boldsymbol{\gamma}$ indexed by latitude (rows) and longitude (columns). 116

A.1 MAP and 95% CrI estimates of the second-order polynomial trend component, $\eta_{1,1}$, indexed by latitude (rows) and longitude (columns). 141
A.2 MAP and 95% CrI estimates of the second-order polynomial trend component, η_2, indexed by latitude (rows) and longitude (columns). 141
A.3 MAP and 95% CrI estimates of the annual amplitude, A_1^j, indexed by latitude (rows) and longitude (columns). 142
A.4 MAP and 95% CrI estimates of the semi-annual amplitude, A_2^j, indexed by latitude (rows) and longitude (columns). 142

A.5 MAP and 95% CrI estimates of the quarterly amplitude, A_4^j, indexed by latitude (rows) and longitude (columns). 142
A.6 MAP and 95% CrI estimates of the annual phase, P_1^j, indexed by latitude (rows) and longitude (columns). 143
A.7 MAP and 95% CrI estimates of the semi-annual phase, P_2^j, indexed by latitude (rows) and longitude (columns). 143
A.8 MAP and 95% CrI estimates of the quarterly phase, P_4^j, indexed by latitude (rows) and longitude (columns). 143

Abstract

Flexible Dynamic Quantile Linear Models

by

Raquel A. Barata

Motivated by the problem of modeling time-varying quantiles in a way that provides rich quantitative information, we consider a class of models to describe the dynamics of a specific quantile for both univariate and multivariate time series data. This prompts us to present several methodological and computational contributions to dynamic quantile modeling, and, more generally, non-Gaussian time-varying models.

We begin with a discussion of the existing quantile estimation literature and the scope to which our methods contribute to the statistical community. We also discuss background on methods for atmospheric river characterization, an application that in part motivated this work and resurfaces throughout.

In the second chapter, we develop a flexible dynamic quantile linear model (exDQLM) utilizing a recently developed family of parametric distributions for quantile regression. A simulation study illustrates our exDQLM to be more robust than the standard Bayesian parametric quantile regression approach for non-standard distributions, performing better in both quantile estimation and predictive accuracy. In addition to a Markov chain Monte Carlo (MCMC) algorithm, we develop an efficient importance sampling variational Bayes (ISVB) algorithm for fast approximate Bayesian inference which is found to produce comparable results to the MCMC in a fraction of the computation time.

In the third chapter, we apply the exDQLM to the analysis of the integrated water vapor transport (IVT) magnitude quantile threshold, a primary component

of many atmospheric river detection schemes. In contrast to current estimation methods, our methodology enables versatile, structured, and informative estimation of the threshold. Further, we develop a transfer function extension to our exDQLM as a method for quantifying non-linear relationships between a quantile of a climatological response and an input. The utility of our transfer function exDQLM is demonstrated in capturing both the immediate and lagged effects of El Niño Southern Oscillation Longitude Index on the estimation of the 0.85 quantile IVT.

In the fourth chapter, we present the R package **exdqlm** as a tool for dynamic quantile regression. The main focus of the package is to provide a framework for Bayesian inference and forecasting of exDQLMs by implementing the methods detailed in the previous two chapters. Non-time-varying quantile regression models, which comprise a majority of the current statistical software, are discussed as a special case of our methods. The software provides the choice of two different algorithms, MCMC or ISVB, for posterior inference. Routines for estimation of a nonlinear relationship via a transfer function model are available as well as routines for forecasting and model evaluation. We illustrate the implementation of the functions and algorithms in the **exdqlm** package with a step-by-step guide for the analysis of several real data sets.

In the fifth chapter, we develop a multivariate extended dynamic quantile linear model to consider multiple time series simultaneously and jointly estimate a specified quantile for each series. To do this, we first develop a multivariate exAL distribution. We then present the details of multivariate MCMC and ISVB algorithms for exact and approximate posterior inference, respectively. The utility of the multivariate model is illustrated via application to two real datasets, including an IVT dataset spanning all of CA.

Finally, we conclude with a brief review of the methodological and computational contributions presented, and discuss possible future work.

Acknowledgments

I was exceptionally lucky work with both Bruno Sansó and Raquel Prado. Their guidance, patience, and support were pivotal to my success. Their fluency in statistics will forever impress and inspire me.

I am grateful for the UCSC statistics faculty as a whole; for their excellent teaching, mentorship, and sense of humor. Also to the students I became friends with along the way; their camaraderie contributed vastly to my enjoyment of graduate school.

My husband Jared, our son Julian, and all our fur children have provided me with distraction and perspective throughout the years which has helped me maintain a healthy balance between my studies and life outside research.

I would like to thank Bin Guan and Duane Waliser at NASA Jet Propulsion Laboratory for sharing their IVT and AR datasets, and being so responsive to any questions I had. The AR database is available at https://ucla.app.box.com/v/ARcatalog and the ERA5 dataset is avaliable at https://cds.climate.copernicus.eu. Also, Christina Patricola at Lawrence Berkeley National Laboratory for helpful conversation about the ELI. The ELI is available at https://portal.nersc.gov/archive/home/projects/cascade/www/ELI. And finally, Shawn Chartrand for sharing his Big Tree Water flow dataset.

The contents of Chapters 2 and 3 are an expansion on the article:

> Raquel Barata, Raquel Prado, and Bruno Sansó. Fast inference for time-varying quantiles via flexible dynamic models with application to the characterization of atmospheric rivers. *The Annals of Applied Statistics*, in press.

I thank the editors and reviewers of *The Annals of Applied Statistics*, whose comments and suggestion strengthened these chapters.

The contents of Chapters 4 are an adaptation of the **exdqlm** R package vignette which has been submitted for publication:

Raquel Barata, Raquel Prado, and Bruno Sansó. exdqlm: An R Package for Estimation and Analysis of Flexible Dynamic Quantile Linear Models. *Journal of Statistical Software*, submitted.

Chapter 1

Introduction

Quantile estimation has been established as a robust alternative to traditional mean-centric approaches. In general, quantiles provide more comprehensive information about the shape of a distribution, which is of particular importance in climatological and environmental applications where changes in variability and extreme events are expect. Nevertheless, methods for quantile estimation in the dynamic setting remain limited as estimation of non-Gaussian time-evolving quantiles is a non-trival task. This work combines well-known quantile regression and dynamic modeling techniques to develop several methodological and computational contributions for dynamic quantile modeling, and, more generally, non-Gaussian time-varying models.

This work is in part motivated by the need to describe and quantify atmospheric rivers (ARs) in global climate and weather models. ARs are elongated regions of water vapor in the atmosphere that transport water from the tropics around the globe. Typically carrying 7-15 Mississippi Rivers worth of water, ARs play a key role in the global water cycle and regional weather (Ralph and Dettinger, 2011). When these events make landfall, they often release water vapor in the form of rain or snow. While they have the ability to cause extreme rainfall

and floods, atmospheric rivers are crucial to alleviate drought conditions and contribute to beneficial increases in the fresh water supply for water-stressed areas such as California (Guan et al., 2010; Dettinger, 2013). Several techniques have been developed with the objective of detecting ARs (Rutz et al., 2014; Backes et al., 2015). In particular, an effective approach is to focus on the integrated water vapor transport (IVT), a vector representing the total amount of water vapor being transported in an atmospheric column. This is increasingly used in the study of ARs because of its direct relationship with orographically induced precipitation (Neiman et al., 2009). One study in particular by Guan and Waliser (2015) presents a method for detection of ARs based on characteristics of the magnitude of the IVT vector. A key component of this and many other AR detection schemes is the thresholding of IVT magnitude at a specified quantile, specifically the 0.85 quantile in Guan and Waliser (2015). A sensitivity study found their AR detection scheme to be sensitive to the choice of quantile, thus accurate estimation of IVT threshold is crucial. However, the current approach for calculating the 0.85 quantile is unstructured, invariant from year to year, unable to incorporate information between locations, and incapable of including relevant climatological information. Motivated by the problem of modeling time-varying IVT thresholds in a way that provides richer quantitative information, we consider a class of models to describe the dynamics of a specific quantile of both univariate and multivariate time series data.

The first contribution of this work is a novel model referred to as the extended dynamic quantile linear model (exDQLM). Our exDQLM utilizes a recently developed family of parametric error distributions for quantile regression, the extended asymmetric Laplace distribution (exAL; Yan and Kottas, 2017). In the Bayesian setting, parametric quantile regression models are almost exclusively based on the

asymmetric Laplace (AL) distribution, a special case of the exAL. However the AL is known to have several drawbacks, which we discuss in detail in Section 2.1. For example, the skewness of the distribution as well as the location of the mode are fully dictated by the choice of the fixed quantile. More flexible error distributions for a single quantile have been considered extensively in the Bayesian non-parametric literature. The median regression case has been considered in the semi-parametric setting by Walker and Mallick (1999), Kottas and Gelfand (2001) and Hanson and Johnson (2002), with general quantile regression seen in Kottas and Krnjajić (2009) and Reich et al. (2009). Fully non-parametric nonlinear modeling of a single quantile regression functions is seen in Taddy and Kottas (2010) and Kottas and Krnjajić (2009). The literature on parametric approaches that lead to flexible quantile regression models is much less extensive. Wichitaksorn et al. (2014) presents a new class of skew distributions with the AL as a special case, however the skewness remains fully determined by the fixed quantile. Zhu and Zinde-Walsh (2009) and Zhu and Galbraith (2011) present a four parameter family of asymmetric exponential power distributions for a fixed quantile, however, the mode of the distribution remains fixed at the quantile of interest. The exAL presented in Yan and Kottas (2017) overcomes these shortcomings in the current parametric methods. A detailed discussion of the properties of the exAL can be found Section 2.1. Our methods generalize the utility of the exAL to the time series setting and allow for time-varying quantile inference.

It is important to mention our exDQLM is not exempt to the possibility of quantile crossing for settings in which multiple quantiles are of interest. This is a well known challenge for the majority of models which provide inference for a single quantile at a time. Nonparametric methods for simultaneous analysis of several quantiles can be found in Reich and Smith (2013) and Tokdar et al.

(2012), however these nonparametric methods are computationally taxing and do not scale well to the time-varying setting. In the context of our application, quantile crossing is not a concern as we are interested in a single fixed quantile.

The second contribution of this work is our importance sampling variational Bayes (ISVB) algorithm for fast, flexible inference of a time-varying quantile. Current methods for quantile regression with time-evolving parameters in both the parametric and semi-parametric approaches are almost exclusively based on the AL likelihood and check loss function, respectively (Gonçalves et al., 2020; Bernardi et al., 2016; Paraschiv et al., 2016; Koenker and Xiao, 2006). Nonparametric approaches are even more limited in the time series setting as defining likelihood functions for quantile-function-valued data is a non-trivial task (Chen et al., 2017). Further, a majority of these approaches, both parametric and nonparametric, are computationally expensive. This has prompted the development of approximate estimation algorithms. Although these alternative algorithms are faster computationally, many compromise the true underlying estimation problem in their original models. Our ISVB algorithm relieves the computational burden while preserving the underlying model structure, thus not compromising the interpretability of the resulting estimated quantiles.

Although variational methods are becoming more widespread in the statistical community (for a comprehensive review see Blei et al. (2017)), the literature on variational inference in the time-varying setting remains limited. There has been some work on linear Gaussian state-space models (Barber and Chiappa, 2007; Penny et al., 2003), dynamic generalized linear models (Quiroz et al., 2018; Berry and West, 2020), and hidden Markov models (Johnson and Willsky, 2014; Foti et al., 2014), however the application of variational inference to non-conjugate nonlinear state-space models is not seen in the literature. Further, the parameteric

quantile regression methods found in Kozumi and Kobayashi (2011) and Yan and Kottas (2017) are special cases of our exDQLM, thus our ISVB can be applied to these non-time-varying models as well. Until this point, variational inference has not been applied in the quantile regression setting. Our ISVB algorithm contributes to both the quantile regression and time-series literature a simple and straightforward method for estimation of the true high-dimensional posterior distribution.

The third contribution of this work is the development of a transfer function exDQLM as a method for quantifying associations that account for the cumulative effect of a time-varying input on a quantile of a response variable, e.g., a given climatological response. Most studies associating climate indices to specific atmospheric phenomena focus on simple linear associations, when in reality the relationships are much more complex. Numerous climate indices have been extensively studied as potential sources of predictability for precipitation and ARs. A few examples include the Arctic Oscillation (AO) index (Guan et al., 2013), the "Pineapple Express" (PE) index (Weller et al., 2012), the Madden-Julian Oscillation (MJO) (Guan et al., 2012), the Niño3.4 index (Tziperman et al., 1998), as well as the recently developed El Niño Southern Oscillation (ENSO) Longitude Index (ELI; Williams and Patricola, 2018). In this work, we demonstrate the practical utility of our transfer function exDQLM in capturing both immediate and lagged effects of ELI on the 0.85 quantile IVT magnitude.

The fourth contribution of this work is the R (R Core Team, 2013) package **exdqlm** which implements the contributions discussed thus far. Although quantile regression has gained popularity, software implementations remain limited, particularly for time-varying methods. Well-developed software implementations of a non-time-varying quantile regression models which utilize the AL are available in

the Bayesian R package **bayesQR** (Benoit et al., 2011) and, equivalently, a classical version, which utilizes the check-loss function (Koenker, 2005), is available in the package **quantreg** (Koenker, 2021). As a result of the AL being a special case of the exAL, the methods used in the Bayesian package **bayesQR** are a special case of the exDQLM implemented in **exdqlm**. Further, there are many established packages for dynamic modeling including **dlm** (Petris and Gilks, 2018) and **dynr** (Ou et al., 2021), however, implementations of dynamic quantile modeling are much more limited. To our knowledge, the package **quantreg** includes a routine which allows for a time-evolving structure in the regression coefficients (a special case of the routines in the package), but this is the extent of the software available for dynamic quantile regression. Our **exdqlm** package provides a powerful and versatile tool for dynamic quantile modeling.

The final contribution of this work is the development of the multivariate exDQLM (MVexDQLM). This model considers multiple time series simultaneously and jointly estimates a specified quantile for each series. Non-time-varying multivariate and spatial quantile regression has been considered in this literature non-parametrically (Reich et al., 2011) and parametrically (Lum et al., 2012). Some work has also been done in the spatio-temporal setting, both parametric (Neelon et al., 2015) and semi-paramteric (Reich, 2012), however again, the parametric approaches are exclusively based on the AL. We develop a multivariate exAL distribution (MVexAL) to facilitate the extension of our more flexible methods to the multivariate time-varying setting. Our methodology, including our variational approach for fast, accurate, approximate estimation, naturally scales to the higher dimension, thus providing a computationally feasible method for estimating conditional quantiles of multivariate time-series.

The details of these topics are presented in the following chapters, organized

as follows. In Chapter 2, we begin with background on the exAL distribution and develop the exDQLM. We present both a Markov chain Monte Carlo (MCMC) algorithm and the ISVB algorithm for Bayesian inference. In Chapter 3, we develop our transfer function exDQLM with details on MCMC and ISVB algorithm augmentations for this new model. We apply the transfer function exDQLM to the estimation of the 0.85 quantile IVT threshold. In Chapter 4, we present the capabilities of the R package **exdqlm** and illustrate its implementation on several real datasets. In Chapter 5, we develop the MVexDQLM as well as the corresponding MCMC and ISVB algorithms. The utility of MVexDQLM is illustrated via application to two real datasets, including an IVT dataset spanning all of CA. Lastly, we conclude in Chapter 6 with a brief review of the methods covered, and discuss possible future work.

Chapter 2

Fast Inference for Time-Varying Quantiles via Flexible Dynamic Models

We begin by developing a flexible dynamic quantile linear model utilizing a recently developed family of parametric distributions for quantile regression. In addition to a Markov chain Monte Carlo algorithm, we develop an efficient variational Bayes algorithm for fast approximate Bayesian inference. A simulation study illustrates the advantages of the exDQLM over the standard Bayesian parametric quantile regression approach, performing better in both quantile estimation and predictive accuracy.

2.1 Introduction and Background

As mentioned previously, Bayesian parametric quantile regression models are almost exclusively based around the asymmetric Laplace (AL) likelihood,

$$\mathrm{AL}_p(y|\mu,\sigma) = \frac{p(1-p)}{\sigma}\exp\left\{-\frac{\rho_p(y-\mu)}{\sigma}\right\} \tag{2.1}$$

where $\rho_p(u) = u[p - I(u < 0)]$ is the check loss function and $I(\cdot)$ denotes the indicator function. $\sigma > 0$ is a scale parameter, $p \in (0,1)$ is a skewness parameter typically fixed to be the quantile of interest, and the mode μ is the corresponding value of that p-th quantile. More explicitly, $\int_{-\infty}^{\mu} \mathrm{AL}_p(y|\mu,\sigma)\mathrm{d}y = p$. A model for quantile regression can be developed by allowing μ to be a function of covariates $\mathbf{x}$, such as $\mu = \mathbf{x}^T\boldsymbol{\beta}^p$ which yields a linear quantile regression structure. Maximization of the AL likelihood with respect to $\boldsymbol{\beta}^p$ is equivalent to the minimization of the check loss function, a common approach in classical quantile regression (Koenker, 2005). For a time-evolving y_t, a time-evolving mode $\mu_t = \mathbf{F}_t'\boldsymbol{\theta}_t^p$ yields a dynamic linear regression structure where $\mathbf{F}_t$ is the regression vector of the covariates corresponding to the parameter vector $\boldsymbol{\theta}_t^p$ at time t. In quantile regression, the parameter vectors are dependent on the fixed quantile p, however for notational simplicity we will omit the superscript p going forward.

The AL was first used for Bayesian quantile regression by Yu and Moyeed (2001) and Tsionas (2003). Kotz et al. (2001) presents several representations of the AL, one of which is a location-scale mixture which easily facilitates posterior simulation (Kozumi and Kobayashi, 2011). That is,

$$\mathrm{AL}_p(y|\mu,\sigma) = \int_{\mathbb{R}^+} \mathrm{N}(y|\mu + A(p)v, \sigma B(p)v)\mathrm{Exp}(v|\sigma)\mathrm{d}v \tag{2.2}$$

where $A(p) = \frac{1-2p}{p(1-p)}$, $B(p) = \frac{2}{p(1-p)}$ and $\mathrm{Exp}(v|\sigma)$ denotes the exponential distri-

bution with mean σ. Although the representation enables closed form posterior conditional distributions, the AL is known to have several limitations. Most notably, the skewness and quantile are fully dictated by choice of p, thus for a fixed quantile the skewness of the distribution is fully determined. In particular, when $p = 0.5$ the distribution is symmetric. Further, for any quantile, the mode of the distribution occurs at μ resulting in rigid tails for extreme percentiles.

To address the shortcomings of the AL parametrically, Yan and Kottas (2017) develop an extension of the AL which overcomes the restrictive aspects of the distribution. The new family of error distributions is constructed from an extension of the location-scale mixture representation of the AL in Equation (2.2). More specifically, replacing the Gaussian kernel in the mixture with a skew-normal distribution introduces an additional skewness parameter γ. When $\gamma = 0$, the model reduces to the AL. The skew-normal density can also be written as a location normal mixture with mixing distribution given by the standard normal truncated to the positive real numbers, facilitating posterior simulation (Henze, 1986). Thus, the full mixture representation of the proposed family of error densities, $\text{exAL}(y|\mu, \sigma, \gamma)$, is

$$\int\int_{\mathbb{R}^+\times\mathbb{R}^+} \text{N}(y|\mu + C(p,\gamma)\sigma|\gamma|s + A(p)v, \sigma B(p)v)\text{Exp}(v|\sigma)\text{N}^+(s|0,1)\text{d}v\text{d}s \quad (2.3)$$

where $\text{N}^+(s|0,1)$ denotes a normal distribution truncated to the positive reals with mean 0 and variance 1. Note that in this parameterization μ no longer corresponds to the p-th quantile of the distribution. To preserve the ability to fix the quantile of interest, which we will now denote to as p_0, Yan and Kottas (2017) define the previously fixed parameter p such that $p = p(p_0, \gamma) = I(\gamma < 0) + \{[p_0 - I(\gamma < 0)]/g(\gamma)\}$ where $g(\gamma) = 2\Phi(-|\gamma|)\exp(\gamma^2/2)$ and $\Phi(\cdot)$ denotes the standard normal CDF. The parameter γ has bounded support over the interval

(L, U) where L is the negative root of $g(\gamma) = 1 - p_0$ and U is the positive root of $g(\gamma) = p_0$. Further, $A(p)$ and $B(p)$ are functions of p as defined in Equation (2.2) and $C(p, \gamma) = [I(\gamma > 0) - p]^{-1}$. Thus, by construction μ corresponds to the fixed quantile p_0 such that $\int_{-\infty}^{\mu} \text{exAL}(y|\mu, \sigma, \gamma)\text{d}y = p_0$.

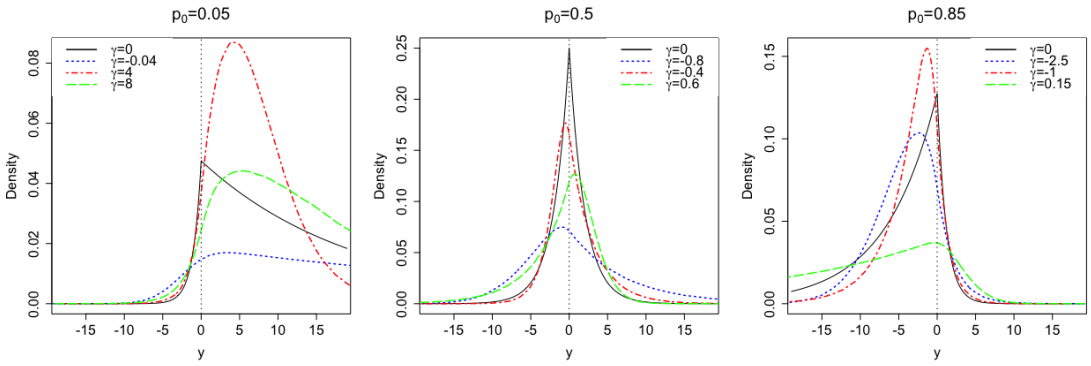

Figure 2.1: Density functions of $\text{exAL}_{p_0}(y|\gamma, \mu, \sigma)$ defined in Equation (2.3) with $\mu = 0$, $\sigma = 1$ and different values of γ for fixed quantiles $p_0 = 0.05$, 0.5, and 0.85. The black solid line corresponds to the AL density, which is a special case when $\gamma = 0$.

Figure 2.1 illustrates the flexibility induced by the additional skewness parameter γ for fixed quantiles $p_0 = 0.05$, 0.5, and 0.85. Recall γ has bounded support on the interval (L, U) which is dependent on γ, thus $\gamma = 0$ is the only skewness parameter which appears in all fixed quantiles of Figure 2.1. It can be seen that, when the median is fixed, γ enables both left and right skewness. The additional parameter controls the tail behavior allowing both heavier and lighter tails than the AL. Flexibility is also seen in the mode, which is no longer fixed at μ. Thus, the exAL is substantially more versatile than the AL while the hierarchical mixture representation preserves straight-forward posterior inference, making it a robust error distribution for our dynamic quantile model. While the exAL improves upon current methods which are well known to suffer from model misspecification (Komunjer, 2005), it remains important that we caution the reader of utilizing the exDQLM beyond quantile estimation. For closed form representa-

tion as well as other properties of the new family of error distributions, see Yan and Kottas (2017). Note also that Yan and Kottas (2017) refers to the extension as the generalized asymmetric Laplace distribution, however we will refer to the distribution as the extended AL (exAL) to avoid confusion with the generalized asymmetric Laplace distribution defined in Kotz et al. (2001).

2.2 The exDQLM

Consider a set of time-evolving responses, y_t, for times $t = 1, \ldots, T$. For each t, a general dynamic model can be defined by

$$\text{Observation equation:} \qquad y_t = \mathbf{F}_t'\boldsymbol{\theta}_t + \epsilon_t \tag{2.4}$$

$$\text{System Equation:} \qquad \boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t. \tag{2.5}$$

Here $\mathbf{F}_t$ is the $q \times 1$ regression vector of the covariates corresponding to the $q \times 1$ regression parameter vector $\boldsymbol{\theta}_t$ at time t, and $\mathbf{G}_t$ is the q-dimensional evolution matrix defining the structure of the parameter vector evolution in time. We propose an extended dynamic quantile linear model (exDQLM) for inference on a single p_0-th quantile by specifying the observational errors of a dynamic linear model to be distributed from the exAL, which we denote exAL_{p_0}. That is, ϵ_t in Equation (2.4) are distributed independently from the exAL with quantile p_0 fixed such that $\int_{-\infty}^{0} \text{exAL}_{p_0}(\epsilon_t|0, \sigma, \gamma)\text{d}\epsilon_t = p_0$. Utilizing a Gaussian time-evolving structure on the system error vector, i.e. $\boldsymbol{\omega}_t \sim \text{N}(\mathbf{0}, \mathbf{W}_t)$ where $\mathbf{W}_t$ is the evolution variance matrix, our exDQLM model can be written

$$y_t|\boldsymbol{\theta}_t, \gamma, \sigma \sim \text{exAL}_{p_0}(\mathbf{F}_t'\boldsymbol{\theta}_t, \sigma, \gamma) \tag{2.6}$$

$$\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \mathbf{W}_t \sim \text{N}(\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t) \tag{2.7}$$

where the normal distribution according to which $\boldsymbol{\theta}_t$ evolves is q-variate. The mixture representation of the exAL in Equation (2.3) can be exploited to rewrite the exDQLM as the following hierarchical model for $t = 1, \ldots, T$:

$$y_t|\boldsymbol{\theta}_t, \sigma, \gamma, v_t, s_t \sim \text{N}(y_t|\mathbf{F}_t'\boldsymbol{\theta}_t + C(p,\gamma)\sigma|\gamma|s_t + A(p)v_t, \sigma B(p)v_t) \tag{2.8}$$

$$v_t, s_t|\sigma \sim \text{Exp}(v_t|\sigma)N^+(s_t|0,1) \tag{2.9}$$

$$\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \mathbf{W}_t \sim \text{N}(\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t). \tag{2.10}$$

Here, $A(p)$, $B(p)$, $C(p,\gamma)$ are the functions of p and γ defined with Equation (2.3). A q-variate prior $\boldsymbol{\theta}_0 \sim \text{N}(\mathbf{m}_0, \mathbf{C}_0)$ is used at the initial stage. It is possible to place an inverse Wishart prior on the evolution covariance matrix $\mathbf{W}_t$, however for our analyses we utilize discount factors, which we discuss in Section 2.2.4. Yan and Kottas (2017) suggest a inverse-gamma prior for σ denoted $\text{IG}(a_\sigma, b_\sigma)$ and uniform prior for γ over the interval (L, U) denoted $\text{Uni}(L, U)$. Further discussion of the prior selection and posterior inference of σ and γ can be found in Section 2.2.5.

2.2.1 Markov chain Monte Carlo algorithm

The construction of the exAL through a structured mixture of normal distributions facilitates Bayesian posterior simulation using Markov Chain Monte Carlo (MCMC) with a Metropolis-Hastings (MH) step for the skewness parameter γ. Conditional on the latent variables $\mathbf{v} = \{v_1, \ldots, v_T\}$ and $\mathbf{s} = \{s_1, \ldots, s_T\}$, scale parameter σ and skewness parameter γ, the dynamic regression coefficients can be sampled using a forward filtering backwards sampling (FFBS) algorithm (Carter and Kohn, 1994; Frühwirth-Schnatter, 1994). Full details of our FFBS can be found in Appendix A. MCMC posterior simulation is summarized in Algorithm 1.

Note that if a point mass prior at zero is used for skewness parameter γ, the

Algorithm 1: exDQLM MCMC

Initialize $\sigma^{(0)}, \gamma^{(0)}, \mathbf{v}^{(0)}, \mathbf{s}^{(0)}, \boldsymbol{\theta}_{1:T}^{(0)}$;

for *i = 0, . . . , I-1* **do**

1. Sample $\sigma^{(i+1)}|\boldsymbol{\theta}_{1:T}^{(i)}, \mathbf{v}^{(i)}, \mathbf{s}^{(i)}, \gamma^{(i)}$ from a generalized inverse Gaussian, denoted $\text{GIG}(\lambda_\sigma, \chi_\sigma, \psi_\sigma)$ where

$$\lambda_\sigma = -(a_\sigma + 1.5T), \quad \chi_\sigma = 2b_\sigma + 2\sum_{t=1}^{T} v_t^{(i)} + \sum_{t=1}^{T} \frac{(y_t - \mathbf{F}_t'\boldsymbol{\theta}_t^{(i)} - A(p)^{(i)} v_t^{(i)})^2}{B(p)^{(i)} v_t^{(i)}},$$

$$\psi_\sigma = \sum_{t=1}^{T} \frac{(C(p)^{(i)}|\gamma^{(i)}|s_t^{(i)})^2}{B(p)^{(i)} v_t^{(i)}}.$$

2. Sample $\gamma^{(i+1)}|\boldsymbol{\theta}_t^{(i)}, \mathbf{v}^{(i)}, \mathbf{s}^{(i)}, \sigma^{(i)}$ using a Metropolis-Hastings step with a Gaussian random walk proposal on the logit scale.

3. **for** *t=1,. . . ,T* **do**

Sample $v_t^{(i+1)}|\boldsymbol{\theta}_t^{(i)}, s_t^{(i)}, \sigma^{(i)}, \gamma^{(i)} \sim \text{GIG}(\lambda_{v_t}, \chi_{v_t}, \psi_{v_t})$ where $\lambda_{v_t} = 1/2$,

$$\chi_{v_t} = \frac{(y_t - \mathbf{F}_t'\boldsymbol{\theta}_t^{(i)} - \sigma C(p)^{(i)}|\gamma^{(i)}|s_t^{(i)}))^2}{\sigma^{(i)} B(p)^{(i)}}, \quad \psi_{v_t} = \frac{2}{\sigma^{(i)}} + \frac{A(p)^{(i)2}}{\sigma^{(i)} B(p)^{(i)}}.$$

end

4. **for** *t=1,. . . ,T* **do**

Sample $s_t^{(i)}|\boldsymbol{\theta}_t^{(i)}, v_t^{(i)}, \sigma^{(i)}, \gamma^{(i)} \sim \text{N}^+(\mu_{s_t}, \sigma_{s_t}^2)$, where

$$\sigma_{s_t}^2 = \left[\frac{C(p)^{(i)2}\gamma^{(i)2}\sigma^{(i)}}{B(p)^{(i)} v_t^{(i)}} + 1\right]^{-1},$$

$$\mu_{s_t} = \sigma_{s_t}^2 \left[\frac{C(p)^{(i)}|\gamma^{(i)}|(y_t - \mathbf{F}_t'\boldsymbol{\theta}_t^{(i)} - A(p)^{(i)} v_t^{(i)})}{B(p)^{(i)} v_t^{(i)}}\right].$$

end

5. **for** *t=1,. . . ,T* **do**

Sample $\boldsymbol{\theta}_t|\mathbf{v}^{(i)}, \mathbf{s}^{(i)}, \gamma^{(i)}, \sigma^{(i)}$ via FFBS. The forward part of the FFBS algorithm uses the forecast distribution $p(y_t|D_{t-1}) = \text{N}(f_t, Q_t)$ where $D_{t-1} = \{y_1, \ldots, y_{t-1}\}$,

$$f_t = \mathbf{F}_t'\mathbf{a}_t + C(p)^{(i)}\sigma^{(i)}|\gamma^{(i)}|s_t^{(i)} + A(p)^{(i)} v_t^{(i)}, \quad Q_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + \sigma^{(i)} B(p)^{(i)} v_t^{(i)}.$$

end

end

model simplifies to the DQLM with fixed quantile $p = p_0$. The DQLM models the p-th quantile alternatively by specifying the observational errors ϵ_t in Equation (2.4) to be distributed independently from an AL (Gonçalves et al., 2020). Similar

to the exDQLM, a mixture representation can be exploited to rewrite the DQLM a hierarchical model to facilitate a MCMC algorithm for posterior inference. Such algorithm will follow very closely Algorithm 1 with a few changes: the posterior of σ reduces to an inverse gamma and all terms with skewness γ will simplify to 0. We compare the exDQLM with this special case in Section 2.3.

2.2.2 Importance sampling variational Bayes algorithm

The addition of two latent parameters per observation in the hierarchical representation of the exDQLM elicits both high computational and memory costs. In particular, any sort of model selection is completely intractable when using only the MCMC algorithm. For example, a daily IVT magnitude time series at a single location from 1979 through 2019 consist of 14965 time points. For this length of time series, the personal laptop used for computations in Chapter 3 was not able to produce results when using the MCMC algorithm due to the significant memory storage required. Further, even with enough memory, model selection with the MCMC algorithm is not a realistic option as it can take a full day or more for the algorithm to converge for each model configuration. Amidst our efforts to address these issues, we found many of the standard methods for fast inference of a non-Gaussian state space model (i.e., an expectation maximization algorithm, or state-space augmentation scheme) were unable to provide accurate inference or compromised the ability to fix the quantile of interest due to the complex structure of our exDQLM. To relieve the computational burden and memory storage requirements induced by the MCMC algorithm while preserving the underlying parameter estimation problem, we present an efficient importance sampling variational Bayes (ISVB) algorithm.

Variational Bayes (VB) is an optimization method for fast, approximate poste-

rior inference (Ostwald et al., 2014). Let $\boldsymbol{\xi} = \{\boldsymbol{\theta}_{1:T}, \sigma, \gamma, \mathbf{v}, \mathbf{s}\}$ denote the set of all parameters in the exDQLM. Within the VB framework, we approximate the posterior distribution $f(\boldsymbol{\xi}|\mathbf{y}_{1:T})$ with an arbitrary variational distribution $r(\boldsymbol{\xi})$ which minimizes the Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951). In particular,

$$r(\boldsymbol{\xi}) := \text{argmin}_{r^*} \text{KL}(r^*(\boldsymbol{\xi}), f(\boldsymbol{\xi}|\mathbf{y}_{1:T})) \tag{2.11}$$

where

$$\begin{aligned} \text{KL}(r(\boldsymbol{\xi}), f(\boldsymbol{\xi}|\mathbf{y}_{1:T})) &= \int r(\boldsymbol{\xi}) \log \frac{r(\boldsymbol{\xi})}{f(\boldsymbol{\xi}|\mathbf{y}_{1:T})} \text{d}\boldsymbol{\xi} & (2.12) \\ &= \int r(\boldsymbol{\xi}) \log \frac{r(\boldsymbol{\xi}) f(\mathbf{y}_{1:T})}{f(\boldsymbol{\xi}, \mathbf{y}_{1:T})} \text{d}\boldsymbol{\xi} & (2.13) \\ &= f(\mathbf{y}_{1:T}) - \int r(\boldsymbol{\xi}) \log \frac{f(\boldsymbol{\xi}, \mathbf{y}_{1:T})}{r(\boldsymbol{\xi})} \text{d}\boldsymbol{\xi}. & (2.14) \end{aligned}$$

This form of the KL divergence leads to the following composition of the log marginal likelihood, or log evidence, which is at the core of the VB approach (Ostwald et al., 2014),

$$f(\mathbf{y}_{1:T}) = \int r(\boldsymbol{\xi}) \log \frac{f(\boldsymbol{\xi}, \mathbf{y}_{1:T})}{r(\boldsymbol{\xi})} \text{d}\boldsymbol{\xi} + \text{KL}(r(\boldsymbol{\xi}), f(\boldsymbol{\xi}|\mathbf{y}_{1:T})). \tag{2.15}$$

The first term in this composition is often referred to as the evidence lower bound (ELBO). Because the KL divergence is non-negative, the variation distribution $r(\boldsymbol{\xi})$ which minimizes the KL divergence in Equation (2.11) also maximizes the ELBO, that is

$$r(\boldsymbol{\xi}) := \text{argmax}_{r^*} \text{ELBO}(r^*(\boldsymbol{\xi}), f(\boldsymbol{\xi}|\mathbf{y}_{1:T})). \tag{2.16}$$

For a full review of the VB approach, see Ostwald et al. (2014).

A common choice for the family of variational distributions over which we

optimize the ELBO is a factorization over different sets of variables known as a mean-field approximation (Beal, 2003). In our particular model, we factorize as follows

$$r(\boldsymbol{\xi}) = r(\boldsymbol{\theta}_{1:T})r(\sigma, \gamma)r(\mathbf{v})r(\mathbf{s}). \tag{2.17}$$

Note, this reflects an assumption of stochastic independence between sets of variables. It has been shown that for each component of the factorization, the ELBO is maximized by the following

$$r(\boldsymbol{\xi}_c) \propto \exp\left\{\int \log f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{-c})r(\boldsymbol{\xi}_{-c})\mathrm{d}\boldsymbol{\xi}_{-c}\right\} \tag{2.18}$$

where $\boldsymbol{\xi}_c$ denotes the set of variables in the component being maximized and $\boldsymbol{\xi}_{-c}$ the variables not in that component of the partition (Tuckerman, 2010).

To implement this VB approach, we initialize the partitioned variational distributions seen in Equation (2.17) and iteratively maximize the ELBO using Equation (2.18) until convergence. For the exDQLM, the variational distribution updates at each iteration are recognizable, closed-form distributions with the exception of $r(\sigma, \gamma)$. Therefore, we propose to approximate the update of $r(\sigma, \gamma)$ at each iteration using importance sampling (IS). ISVB posterior inference for the exDQLM is summarized in Algorithm 2. For simplicity, we will use the following short-hand notation where $\boldsymbol{\xi}_c$ and $\boldsymbol{\xi}_{-c}$ are as defined in Equation (2.18)

$$\langle g(\boldsymbol{\xi}_c)\rangle = \int \log g(\boldsymbol{\xi}_c)r(\boldsymbol{\xi}_{-c})\mathrm{d}\boldsymbol{\xi}_{-c}.$$

The resulting closed form integrals as well as complete details of the Forward Filtering Backwards Smoothing (FFBSm) and IS algorithms used to update the variational distributions can be found in Appendix A.

Algorithm 2: exDQLM ISVB

Set $k = 0$ and initialize $r^0(s_t)$, $r^0(v_t)$, $r^0(\boldsymbol{\theta}_t)$ and $r^0(\sigma, \gamma)$;
while *convergence has not been achieved* **do**

1. **for** t=1,...,T **do**
Update $r^{(k+1)}(v_t) = \text{GIG}(\lambda_{v_t}^{(k+1)}, \chi_{v_t}^{(k+1)}, \psi_{v_t}^{(k+1)})$ where $\lambda_{v_t} = 1/2$

$$\chi_{v_t} = \left\langle \frac{1}{\sigma B(p)} \right\rangle^{(k)} (y_t^2 - 2y_t \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle^{(k)} + \langle (\mathbf{F}_t' \boldsymbol{\theta}_t)^2 \rangle^{(k)}) - 2\langle s_t \rangle^{(k+1)} \left\langle \frac{C(p)|\gamma|}{B(p)} \right\rangle^{(k)} (y_t - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle^{(k)}) + \langle s_t^2 \rangle^{(k+1)} \left\langle \frac{C(p)^2 \sigma |\gamma|^2}{B(p)} \right\rangle^{(k)}$$

$$\psi_{v_t} = 2 \left\langle \frac{1}{\sigma} \right\rangle^{(k)} + \left\langle \frac{A(p)^2}{\sigma B(p)} \right\rangle^{(k)}.$$

end

2. **for** t=1,...,T **do**
Update $r^{(k+1)}(s_t) = \text{N}^+(\mu_{s_t}^{(k+1)}, \sigma_{s_t}^{2\ (k+1)})$, where

$$\sigma_{s_t}^{2(k+1)} = \left[\left\langle \frac{C(p)^2 \sigma \gamma^2}{B(p)} \right\rangle^{(k)} \left\langle \frac{1}{v_t} \right\rangle^{(k)} + 1 \right]^{-1}$$

$$\mu_{s_t}^{(k+1)} = \sigma_{s_t}^2 \left[(y_t - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle^{(k)}) \left\langle \frac{1}{v_t} \right\rangle^{(k)} \left\langle \frac{C(p)|\gamma|}{B(p)} \right\rangle^{(k)} - \left\langle \frac{C(p)|\gamma| A(p))}{B(p)} \right\rangle^{(k)} \right].$$

end

3. **for** t=1,...,T **do**
Update the smoothed distribution $r^{(k+1)}(\boldsymbol{\theta}_t) = \text{N}(\mathbf{m}_t^s, \mathbf{C}_t^s)$ using a FFBSm algorithm with forecast distribution $r^{(k+1)}(y_t | D_{t-1}) = \text{N}(f_t, Q_t)$ where $D_{t-1} = \{y_1, \ldots, y_{t-1}\}$,

$$f_t = \mathbf{F}_t' \mathbf{a}_t + \left[\left\langle \frac{C(p)|\gamma|}{B(p)} \right\rangle^{(k)} \langle s_t \rangle^{(k+1)} + \left\langle \frac{A(p)}{\sigma B(p)} \right\rangle^{(k)} \Big/ \left\langle \frac{1}{v_t} \right\rangle^{(k+1)} \right] \Big/ \left\langle \frac{1}{\sigma B(p)} \right\rangle^{(k)}$$

$$Q_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + \left[\left\langle \frac{1}{v_t} \right\rangle^{(k+1)} \left\langle \frac{1}{\sigma B(p)} \right\rangle^{(k)} \right]^{-1}.$$

end

4. Update $r^{(k+1)}(\sigma, \gamma)$ using IS with proposal distributions $t_{(L,U)}(0, 1)$ and $t_{(0,\infty)}(m_\sigma, v_\sigma)$ for γ and σ, respectively, where m_σ and v_σ denote the mean and variance of the prior distribution on σ. Further details of this IS step can be found in Appendix A.

5. Set $k = k + 1$.

end

2.2.3 Comparison criteria

To evaluate the quantile inference and predictive performance of the exDQLM, we define several measures for comparison. Consider first the setting in which we know the true p_0 quantile, μ_t^{true}, for all t. To measure the fit of the quantile estimates, we compute the 95% credible interval (CrI) for the mean check loss (MCL),

$$\sum_t \rho_{p_0}(\mu_t^{true} - \mathbf{F}_t'\tilde{\boldsymbol{\theta}}_t)/T, \tag{2.19}$$

where $\tilde{\boldsymbol{\theta}}_t$ is a sample from the posterior distribution.

To evaluate the predictive ability of the exDQLM, we consider the Gelfand and Ghosh (1998) posterior predictive loss criterion (pplc) with check loss function ρ_{p_0}. Given the posterior replicate distribution of y_t, $p(y_t^{rep}|D_T)$,

$$\text{pplc} = \sum_t \mathbb{E}[\rho_{p_0}(y_t^{obs} - y_t^{rep})|D_T] \tag{2.20}$$

where $D_T = \{y_1, \ldots, y_T\}$.

Lastly, as in Huerta et al. (2003) and Prado et al. (2006) we use the one-step-ahead predictive distribution function introduced by Rosenblatt (1952) as a model diagnostic tool. If we define $\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}} = \{\mathbf{v}, \mathbf{s}, \sigma, \gamma\}$, this distribution is given by

$$u_t = \Phi(y_t|D_{t-1}, \boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}) = \Pr(Y_t \leq y_t|D_{t-1}, \boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}). \tag{2.21}$$

Here u_t defines an independent sequence which is uniformly distributed on the interval $(0, 1)$ (Rosenblatt, 1952). Conditional on $\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}$, the predictive distribution of y_t is normally distributed with mean f_t and variance Q_t seen in Algorithms 1 and 2, thus $u_t = \Phi(y_t|f_t, Q_t)$ where Φ denotes the normal CDF. We can obtain

a point estimate for u_t conditionally on a posterior summary of $\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}$ from the MCMC or ISVB posterior samples. A diagnosis of the model performance can be done through the correlation of the estimated sequence $\{\hat{u}_t\}$ and their distribution shape. More specifically, transforming the values with a standard normal inverse CDF allows for examination of the distribution shape with a normal QQ-plot. To quantify the divergence from the standard normal distribution, we consider the KL divergence $\mathrm{KL}(h, \phi) = \int_{-\infty}^{\infty} h(x) \log \frac{h(x)}{\phi(x)} \mathrm{d}x$. We estimate the integrals using the numerically approximated densities of our transformed sample, which we denote h, and the standard normal density, ϕ.

2.2.4 Discount factor selection

A standard approach which allows us to specify the time-evolving covariance matrices $\mathbf{W}_t$ is the use of discount factors. (West and Harrison, 2006). The structure and magnitude of $\mathbf{W}_t$ controls stochastic variation and stability of the evolution of the model over time. More precisely, if the posterior variance of the state vector $\boldsymbol{\theta}_{t-1}$ at time $t-1$ is denoted as $\mathrm{Var}(\boldsymbol{\theta}_{t-1}|D_{t-1}) = \mathbf{C}_{t-1}$, the sequential updating equations produce the prior variance of $\boldsymbol{\theta}_t$, $\mathbf{R}_t = \mathrm{Var}(\boldsymbol{\theta}_t|D_{t-1}) = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t' + \mathbf{W}_t$. Between observations, the addition of the error leads to an additive increase in the initial uncertainty $\mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t'$ of the system variance. Thus it is natural to write $\mathbf{R}_t$ as a fixed proportion of $\mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t'$ such that $\mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t'/\delta \geq \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t'$. Here δ is defined to be a discount factor such that $0 < \delta \leq 1$. This suggests an evolution variance matrix of the form $\mathbf{W}_t = \frac{1-\delta}{\delta}\mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t'$, where the $\delta = 1$ results in a static model with non time-varying parameters.

Selection of discount factors is typically done by optimizing some model checking criterion. This criterion-based selection approach requires posterior inference

for each set of discount factors which can become computationally expensive very quickly especially for large T. The ISVB algorithm makes this criterion-based selection approach computationally feasible. We propose selecting the discount factor, or combination of discount factors (see West and Harrison (2006) for details on component discounting), that minimize the KL divergence calculated from the one-step-ahead predictive distribution functions u_t estimated using the MAP estimates of $\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}$ from the ISVB algorithm, as discussed in Section 2.2.3. Fixing the discount factors within each quantile ensures consistent signal-to-noise ratios between differing models and algorithms.

2.2.5 Comments on prior selection and inference of σ and γ

We find that using a weakly informative prior distribution on the skewness parameter γ facilitates reliable posterior inference by alleviating some of the inferential problems known to arise when utilizing the skew-normal family (Liseo and Loperfido, 2006), particularly within the ISVB algorithm. More specifically, for several datasets we find using a flat prior results in the ISVB estimates of γ going to the boundary of its support, resulting in a host of numerical and inferential problems. To this end, we implement a Student-t distribution truncated to the interval (L, U) as the prior for γ, i.e. $\gamma \sim t_{(L,U)}(0, 1)$ with one degree of freedom, in contrast to the uniform prior suggest by Yan and Kottas (2017).

Interaction between the parameters σ and γ can complicate posterior inference, particularly for extreme quantiles. Joint sampling of σ and γ with a random-walk MH step facilitates mixing and convergence within the MCMC algorithm. The interaction between σ and γ is also prevalent within the ISVB algorithm, which commonly results in the variational distributions getting stuck in local optima. To facilitate fast posterior estimation with the ISVB algorithm, we place a point-

mass prior on σ. For the simulation study in Section 2.3, we set the point-mass of this prior to the posterior mode of σ estimated from the DQLM. That is, for any fixed quantile of interest p_0, the prior on σ is set to $\delta_{\hat{\sigma}^{p_0}_{\gamma=0}}(\sigma)$ where δ denotes the Dirac delta function and $\hat{\sigma}^{p_0}_{\gamma=0}$ is the posterior mode of σ under the DQLM for the p_0 quantile. If convergence still proves difficult, a smaller scale can be helpful and we suggest decreasing the location of the point-mass by approximately twenty-five percent, as seen in the IVT analysis of Chapter 3. Although this approach results in different posterior summaries for the skewness parameter γ obtained from the ISVB algorithm and the MCMC algorithm, we find that the posterior error distributions and modes (and therefore quantile estimates) are robust with respect to the prior placed on σ.

2.3 Simulation study

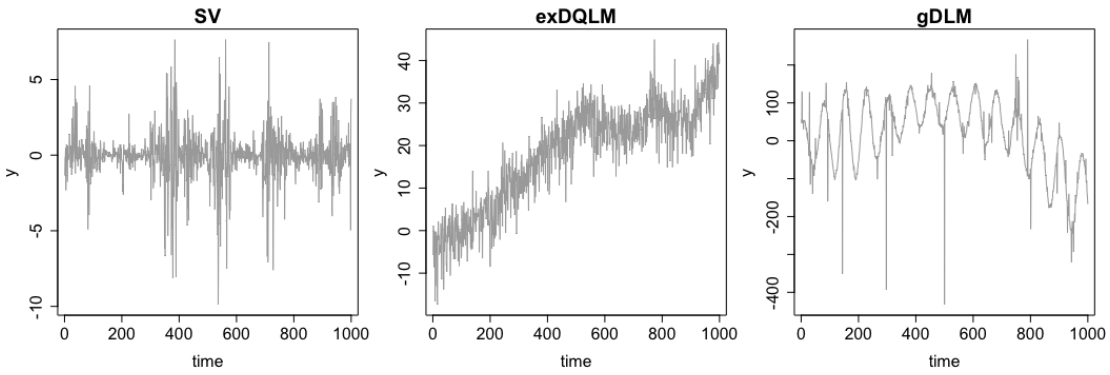

Figure 2.2: Simulation study datasets. From left to right are the data simulated from the Stochastic Volatility (SV) model, exDQLM, and generalized DLM (gDLM) described in Section 2.3.

We present results from a simulation study to examine the exDQLM for three different quantiles; 0.05, 0.50 and 0.85. We compare the flexibility of the model to the special case of the DQLM as well as the standard DLM, for

which we can estimate posterior quantiles from the Gaussian observation equation using the smoothed estimates (West and Harrison, 2006). That is, sample $\mathbf{F}'\tilde{\theta}_t^{DLM} + \tilde{\sigma}^{DLM}\Phi^{-1}(p_0)$ where $\tilde{\theta}_t^{DLM}$ denotes a sample from the smoothed distributions of the DLM state vector, $\tilde{\sigma}^{DLM}$ denotes a sample from the smoothed distribution of the observational standard deviation, and Φ^{-1} denotes the standard normal inverse CDF. For the true underlying data-generating distributions of the simulated datasets, we consider three scenarios with different types of tail behavior and skewness.

Dataset 1: Stochastic Volatility. Stochastic Volatility (SV) models are commonly used to analyze returns (Kastner, 2016). These models are stochastic processes in which the log-variance is randomly distributed and follows an autoregressive structure. The SV model for $t = 1, \ldots, 1000$, where y_t denotes the return at time t, can be written as follows,

$$y_t | x_t \sim \text{N}(0, x_t) \tag{2.22}$$

$$\log x_t | \log x_{t-1}, \mu, \phi, \sigma \sim \text{N}(\mu + \phi\{\log x_{t-1} - \mu\}, \sigma^2) \tag{2.23}$$

$$\log x_0 | \mu, \phi, \sigma \sim \text{N}(\mu, \sigma^2/(1-\phi)) \tag{2.24}$$

Under SV models, posterior inference of the return distributions requires simulation of a latent time-varying process which can sometimes be difficult. To explore the posterior performance of our exDQLM with respect to this more complexly structured data, we generate time series of length 1000 from a SV model using the `stochvol` package in `R` with the level (μ), persistence (ϕ), and volatility (σ) of the log-variance to be 0, 0.95 and 0.5 (Hosszejni and Kastner, 2018), respectively. We will utilize the exDQLM to model the $p_0 = 0.05, 0.5$, and 0.85 posterior quantiles

with a first-order polynomial evolution structure,

$$y_t \sim \text{exAL}_{p_0}(\theta_t, \sigma, \gamma)$$
$$\theta_t \sim \text{N}(\theta_{t-1}, W_t).$$

Dataset 2: exDQLM. Next, we consider synthetic data from an exDQLM, for $t = 1, \ldots, 1000$,

$$y_t \sim \text{exAL}_{0.85}(\mathbf{F}'\boldsymbol{\theta}_t, \sigma, \gamma)$$
$$\boldsymbol{\theta}_t \sim \text{N}_2(\mathbf{G}\boldsymbol{\theta}_{t-1}, \mathbf{W}).$$

With a slight abuse of notation, here $\mathbf{F}'\boldsymbol{\theta}_t$ denotes the $p_0 = 0.85$ quantile of the synthetic dataset at time t. The components $\mathbf{F}$ and $\mathbf{G}$ are specified with a second-order polynomial trend (West and Harrison, 2006),

$$\mathbf{F} = (1, 0)', \quad \mathbf{G} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

with

$$\mathbf{W} = \begin{bmatrix} 0.01 & 0.001 \\ 0.001 & 0.001 \end{bmatrix},$$

$\sigma = 1$, and skewness parameter $\gamma = -2.5$ causing the mode to be below the $p_0 = 0.85$ quantile for all t. We model the $p_0 = 0.05, 0.5, 0.85$ quantiles of this dataset (simulated with p_0 fixed at 0.85) with the second-order polynomial evolutionary structure seen above. We expect to recover the values of σ and γ used to simulate the data for only the estimated 0.85 quantile.

Dataset 3: Generalized DLM. For a dataset with extreme observations, we generate data from a non-Gaussian DLM (West and Harrison, 2006), for $t =$

$1, \ldots, 1000$,

$$y_t \sim \text{Cauchy}(\mathbf{F}'\boldsymbol{\theta}_t, \tau^2)$$
$$\boldsymbol{\theta}_t \sim \text{N}_4(\mathbf{G}\boldsymbol{\theta}_{t-1}, \mathbf{W}).$$

Again, with a slight abuse of notation, here $\mathbf{F}'\boldsymbol{\theta}_t$ denotes the mean of the synthetic dataset at time t. The components $\mathbf{F}$ and $\mathbf{G}$ are specified with a second-order polynomial and Fourier form represented seasonality at frequency $\omega = 2\pi/75$ (West and Harrison, 2006),

$$\mathbf{F} = (1, 0, 1, 0)', \quad \mathbf{G} = \text{block-diag}\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{bmatrix} \right\}$$

with $\tau^2 = 4$ and evolution covariance

$$\mathbf{W} = \text{block-diag}\left\{ \begin{bmatrix} 0.05 & 0.01 \\ 0.01 & 0.001 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \right\}.$$

Again, we model the quantiles with the same trend and seasonal evolution structure.

2.3.1 Results

For all models, we set conjugate prior $\boldsymbol{\theta}_0 \sim \text{N}(\mathbf{m}_0, \mathbf{C}_0)$ and priors for σ and γ as discussed in Section 2.2.5. Table 2.1 reports the posterior results, with bold text indicating the model supported by the comparison criteria detailed in Section 2.2.3.

Overall, the exDQLM out-performs the standard DLM and DQLM. The exDQLM is favored with a lower MCL for all cases in which the true quantile

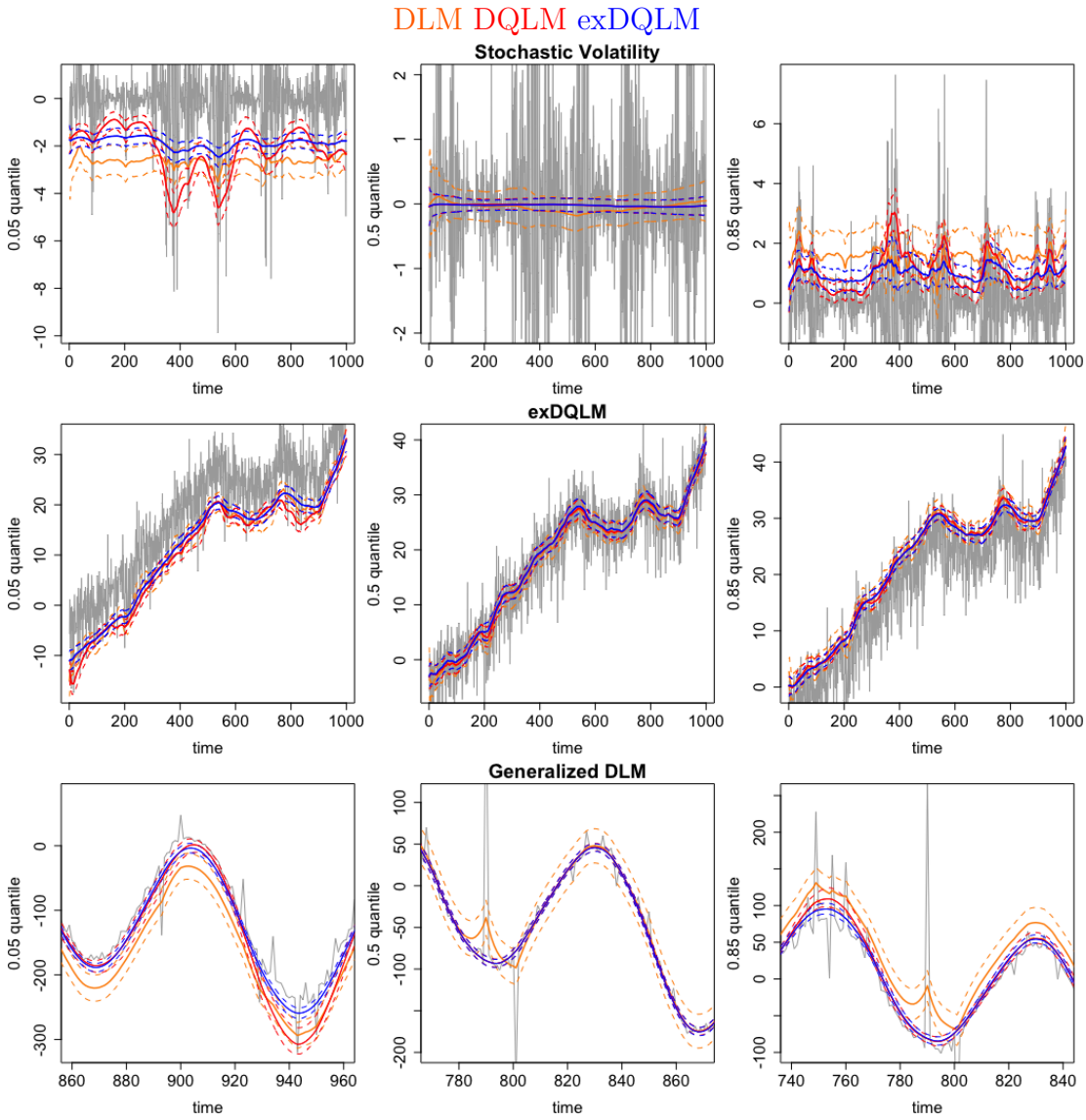

Figure 2.3: MCMC results. Smoothed posterior distributions of dynamic quantiles $p_0 = 0.05, 0.5, 0.85$. exDQLM estimates in blue, DQLM in red, and DLM in orange. Dotted lines indicate the 95% CrI from the smoothed posterior distributions and solid lines indicate posterior mean estimates. Due to the scale of the gDLM data, we focus on time periods which highlight the similarities and/or differences of the models at the three different quantiles.

is known with two exceptions where the MCL of the exDQLM is comparable to the MCL of the DQLM: the medians of the Stochastic Volatility and generalized

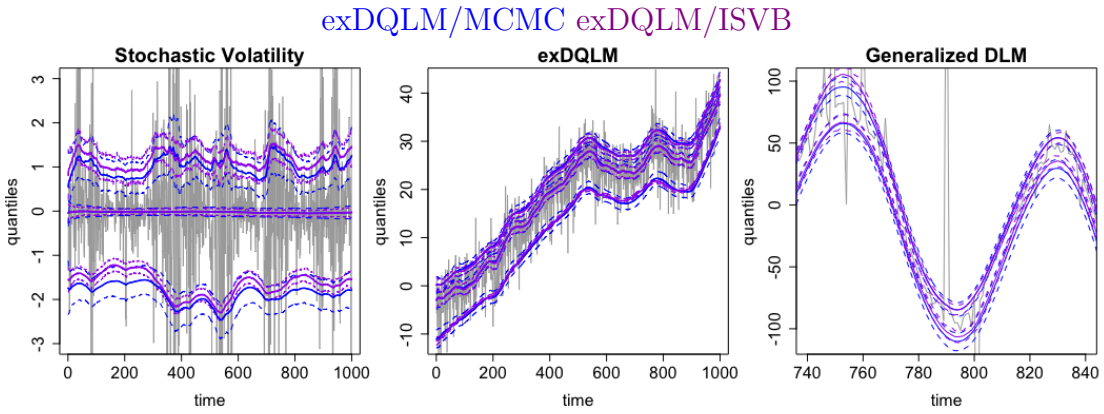

Figure 2.4: ISVB and MCMC comparison results. Posterior distributions of dynamic quantiles $p_0 = 0.05, 0.5, 0.85$. MCMC exDQLM estimates in blue, ISVB exDQLM estimates in purple. Dotted lines indicate the 95% CrI from the smoothed posterior distributions and solid lines indicate posterior mean estimates. Due to the scale of the gDLM data, we focus on period of only 100 time points and omit the median estimates for visual clarity.

DLM, both symmetric datasets. The one-step-ahead predictions assessed using the KL divergence also overwhelmingly favor the exDQLM for all quantiles except 0.5 of the symmetric datasets, in which the exDQLM is again comparable to the DQLM. Similarly, the Gelfand and Ghosh pplc favors the exDQLM for all extreme quantiles (0.05 and 0.85), and again is comparable to the DQLM for the median in the two cases for which the data is symmetric. This parallel between the DQLM and exDQLM for the median of the symmetric datasets is unsurprising, as the exAL reduces to the AL at the 0.5 quantile when the distribution is symmetric. However, we find the pplc for the 0.5 quantile of the exDQLM dataset is also comparable between all three models and even slightly favors the DLM. The exAL distribution used to generate the dataset (seen in Figure 2.1) is only slightly left-skewed around the median with thin tails, therefore it is not unreasonable a normal observational distribution is able to produce equitable predictive results. With this exception, the exDQLM outperforms the other models in all

cases for which the underlying distribution is skewed or the quantile of interest does not align with the mode. Figure 2.3 supports these findings where it can be seen that, due to the rigidity of their observational error distributions, the estimated dynamic quantiles of the DLM and DQLM are more affected by extreme observations than the exDQLM. This is particularly evident for extreme quantiles, i.e., the 0.05 quantile. Our findings highlight the two main advantages of our exDQLM for parametric quantile inference on non-Gaussian dynamic models; robust estimation of any dynamic quantile and superior predictive accuracy for non-standard distributions.

The assessment criteria also illustrate the comparable accuracy of the ISVB exDQLM algorithm to the MCMC exDQLM, but with a fraction of the computational time (see Table 2.1). Both algorithms were implemented in the R programming language on a personal computer with a 2.5 GHz Intel Core i5 processor. The point-mass prior on parameter σ results in different posterior summaries for γ from the MCMC and ISVB algorithms, as discussed in Section 2.2.5 and seen in Table 2.1. It can been seen in Figure 2.4 that the credible intervals (CrIs) of the ISVB algorithm are narrower than the CrIs of the MCMC algorithm. The median proportion of coverage agreement between the 95% CrIs from the ISVB and MCMC algorithms are reported in Table 2.2. It can be seen that the median proportion of the 95% CrIs from the MCMC algorithm contained within the 95% CrIs from the ISVB algorithm varies between approximately 0.33 and 0.83. This underestimation of the variability is a feature to be aware of when using variational methods. However, the approximated posterior quantiles from the ISVB algorithm are almost entirely contained with the MCMC posterior 95% CrIs.

2.4 Conclusion

Motivated by the need for versatile estimation of a single quantile over time, we have presented several methodological and computational contributions for dynamic quantile modeling. Our exDQLM has two main advantages; the model facilitates more flexibility in the estimation of the quantile than standard Bayesian parametric quantile regression approaches, and relevant features such as seasonality or structured long-term variability are easily included in the evolution structure of the quantile. Further, the development of our efficient ISVB algorithm facilitates fast posterior inference, making our methodology accessible even in applications with very long time series data. Our methodology is immediately beneficial not only in climatological applications such as AR detection detailed in the following chapter, but more generally in any application with non-Gaussian time-varying models.

Model	γ	σ	MCL	KL	pplc	time
Stochastic Volatility ($\delta = 0.92, 0.99, 0.87$ for $p_0 = 0.05, 0.50, 0.85$, respectively)						
$p_0 = 0.05$						
DLM	—	—	—	3.644	839.7	0.1
DQLM	—	0.178 (0.17,0.19)	—	3.943	2012.6	19.6
exDQLM/MCMC	4.358 (3.96,4.77)	0.303 (0.29,0.32)	—	**3.641**	**697.5**	46.2
exDQLM/ISVB	6.980 (6.93,7.03)	0.178 (fixed)	—	3.911	774.3	4.1
$p_0 = 0.50$						
DLM	—	—	0.035	3.586	852.6	0.1
DQLM	—	0.486 (0.46,0.52)	**0.011**	**3.484**	**749.7**	21.3
exDQLM/MCMC	0.037 (-0.03,0.09)	0.487 (0.46,0.52)	0.012	3.498	757.4	48.7
exDQLM/ISVB	0.017 (-0.02,0.05)	0.486 (fixed)	0.012	3.524	769.6	0.8
$p_0 = 0.85$						
DLM	—	—	—	3.679	829.5	0.1
DQLM	—	0.299 (0.28,0.32)	—	3.979	1133.3	19.5
exDQLM/MCMC	-1.384 (-1.53,-1.18)	0.401 (0.376,0.423)	—	**3.667**	**740.6**	43.9
exDQLM/ISVB	-2.610 (-2.65,-2.57)	0.299 (fixed)	—	3.652	785.5	2.1
exDQLM (trend $\delta = 0.93$ for all p_0)						
$p_0 = 0.05$						
DLM	—	—	—	4.496	2282.7	0.1
DQLM	—	0.456 (0.43,0.48)	—	3.968	4276.3	18.6
exDQLM/MCMC	5.139 (4.75,5.52)	0.854 (0.81,0.94)	—	**3.910**	**1692.6**	46.2
exDQLM/ISVB	8.058 (8.02,8.09)	0.456 (fixed)	—	4.076	1799.9	4.6
$p_0 = 0.50$						
DLM	—	—	—	4.169	**2287.6**	0.1
DQLM	—	1.584 (1.49,1.68)	—	3.578	2343.8	19.0
exDQLM/MCMC	0.362 (0.27,0.48)	1.377 (1.20,1.53)	—	**3.548**	2328.8	45.9
exDQLM/ISVB	0.25 (0.21,0.28)	1.584 (fixed)	—	3.674	2345.0	0.7
$p_0 = 0.85$						
DLM	—	—	0.214	4.384	2283.3	0.1
DQLM	—	0.871 (0.82,0.93)	0.252	3.939	2963.1	18.9
exDQLM/MCMC	-2.514 (-2.72,-2.39)	0.967 (0.87,1.03)	**0.195**	**3.831**	**1489.7**	46.2
exDQLM/ISVB	-2.68 (-2.71,-2.65)	0.871 (fixed)	0.211	3.835	1510.2	0.7
Generalized DLM (trend $\delta = 0.98$, seasonality $\delta = 0.95$ for all p_0)						
$p_0 = 0.05$						
DLM	—	—	1.830	5.323	84242.3	0.1
DQLM	—	3.164 (2.98,3.37)	1.410	3.729	45607.3	19.0
exDQLM/MCMC	3.469 (3.23,3.76)	3.492 (3.35,3.71)	**0.472**	**3.603**	**10747.9**	45.0
exDQLM/ISVB	1.534 (1.45,1.62)	3.164 (fixed)	1.046	3.609	19189.5	2.4
$p_0 = 0.50$						
DLM	—	—	2.723	5.401	84316.5	0.1
DQLM	—	4.536 (4.26,4.84)	1.204	3.635	7400.3	18.9
exDQLM/MCMC	0.112 (0.06,0.18)	4.465 (4.21,4.71)	**1.203**	**3.581**	**7278.6**	45.5
exDQLM/ISVB	0.095 (0.06,0.13)	4.536 (fixed)	1.240	3.514	7616.1	0.7
$p_0 = 0.85$						
DLM	—	—	4.518	5.339	84347.7	0.1
DQLM	—	3.149 (2.95,3.36)	0.979	3.531	13454.3	18.6
exDQLM/MCMC	-1.139 (-1.24,-1.02)	3.703 (3.478,3.926)	**0.823**	**3.516**	**8873.2**	45.1
exDQLM/ISVB	-0.497 (-0.57,-0.42)	3.149 (fixed)	0.828	3.491	9836.2	1.1

Table 2.1: Posterior summaries for γ and σ (where applicable): mean (95% CrI). Mean check loss of the MAP dynamic quantile. KL divergences of the one-step-ahead distributions. Posterior predictive loss criterion (pplc) under the check loss function. Computation run-time (min).

p_0	SV	exDQLM	gDLM
0.05	0.368	0.327	0.524
0.50	0.709	0.831	0.785
0.85	0.629	0.675	0.685

Table 2.2: Median proportion of coverage agreement between the 95% CrIs from the ISVB and MCMC algorithms. That is, the proportion of the 95% CrIs from the MCMC algorithm covered by the 95% CrIs from the ISVB algorithm.

Chapter 3

Estimating a Non-linear Response in a Time-Varying Quantile with Application to the Characterization of Atmospheric Rivers

In this chapter, we apply the exDQLM to the analysis of the IVT quantile threshold. In doing so, we develop a transfer function extension to our exDQLM as a method for quantifying non-linear relationships between a quantile of a climatological response and an input. The utility of our transfer function exDQLM is demonstrated in capturing both the immediate and lagged effects of El Niño Southern Oscillation Longitude Index on the estimation of the 0.85 quantile IVT.

3.1 Introduction

The method presented in Guan and Waliser (2015) for detection of ARs from the IVT is as follows. For each of the 12 months, the 0.85 quantile IVT is calculated

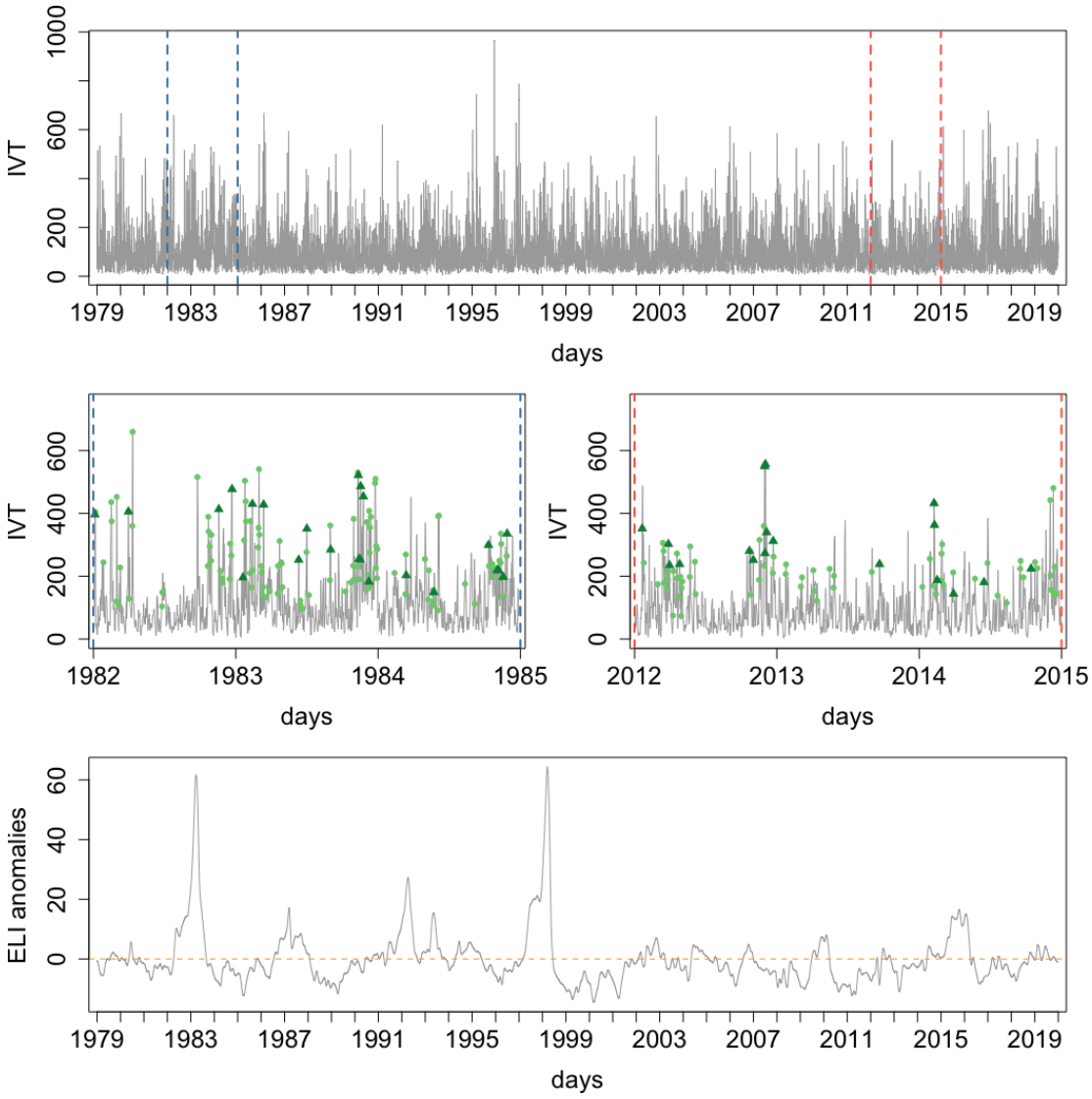

Figure 3.1: Top panel: Average daily IVT magnitude in the grid cell containing Santa Cruz, CA. The dashed vertical lines enclose the two time periods enlarged in the middle panel. Middle panel: A closer look at the two time periods highlighted in the top panel. Years 1982 to 1985 illustrate a time period which saw an exceptional amount of rain. Years 2012 to 2015 illustrate a period which was exceptionally dry. ARs detected by the scheme proposed in Guan and Waliser (2015) in the Santa Cruz grid cell are indicated with dark triangles. ARs detected in neighboring coastal grid cells are indicated with lighter circles. Bottom panel: ELI anomalies resulting from the de-seasonalization of the interpolated ELI. The dashed, horizontal line is at zero, for reference.

over all time steps during the 5 month windows centered on that month over the period from 1997 to 2014 at a specific location. Comparison to the estimated 0.85 quantile in combination with a minimum threshold is used to isolate regions of enhanced IVT as possible ARs. Criteria are then applied to the length and width of these regions, resulting in a defined set of ARs. Finally, the coastal location intersecting with an AR at which the IVT magnitude is highest is defined as the cell in which the AR makes landfall. Guan and Waliser (2015) conduct a sensitivity study in which they consider various possible quantiles, however we focus the following analysis on the 0.85 quantile to illustrate our estimation methods. For more details on the full AR detection algorithm and sensitivity study, see Guan and Waliser (2015).

The primary dataset used to calculate IVT in the study by Guan and Waliser (2015) is the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim reanalysis (ERA-Interim) (Berrisford et al., 2011; Dee et al., 2011), however we utilize the latest ECMWF observational product ERA5 in our analysis. ERA5 produces 1-hourly atmospheric fields at a $0.5° \times 0.5°$ spatial resolution beginning in 1979 and is continuously updated at a 2 month lag (Hersbach et al., 2020). The top panel of Figure 3.1 illustrates the average daily IVT magnitude in Santa Cruz, CA, of which we examine the 0.85 quantile in this analysis. For illustration, the times at which ARs detected to make landfall at that location and in the neighboring coastal locations are illustrated in the middle panels of Figure 3.1 for two time periods; years 1982 to 1985 in which CA saw an exceptional amount of rain, and years 2012 to 2015 which were exceptionally dry for CA.

Although many climate indices other than ELI have been studied as potential sources of predictability for ARs, initial examination of several indices with our transfer function exDQLM did not demonstrate significant associations. There-

fore in this analysis, we focus solely on the association between IVT and ELI. ELI is a single metric which captures the spatial diversity of ENSO, created utilizing the monthly ECMWF twentieth century reanalysis (ERA-20C). In particular, ELI is the average longitude at which tropical Pacific deep convection occurs at a given month. For further details on the development ELI and the relationship of ELI with precipitation see Williams and Patricola (2018) and Patricola et al. (2020), respectively. The monthly ELI dataset is available online beginning in 1854 and is frequently updated (Williams and Patricola, 2018). For our analysis, we interpolate the ELI to the daily time scale and de-seasonalize the time series by removing the smoothed posterior mean estimates from a standard DLM with constant trend, annual, semi-annual, and quarterly components. De-seasonalizing in this way ensures the variability in the 0.85 quantile described by the ELI component of our model is not an artifact of the seasonality in the original ELI time series. This de-seasonalization results in a time series of ELI anomalies, also seen in Figure 3.1, which we use as the input in our analysis of the 0.85 quantile of the IVT magnitude.

Quantifying the relationship between a climatological response and input at various quantiles is a non-trival task. In the mean-centric setting, transfer functions are a simple way to incorporate variables which measure the combined effect of current and past inputs or regression effects (West and Harrison, 2006). To capture both the immediate and lagged effects of a climatological variable, we expand the use of transfer functions to the dynamic quantile regression setting with the development of a transfer function extension to our exDQLM.

3.2 Transfer Function exDQLM

For time $t = 1, \ldots, T$ and a single regression effect, X_t, a transfer function exDQLM with exponential decay is as follows:

$$y_t|\boldsymbol{\theta}_t, \gamma, \sigma \sim \text{exAL}_{p_0}(\mathbf{F}_t'\boldsymbol{\theta}_t + \zeta_t, \sigma, \gamma) \tag{3.1}$$

$$\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \mathbf{W}_t \sim \text{N}(\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t) \tag{3.2}$$

$$\zeta_t|\zeta_{t-1}, \psi_{t-1}, \omega_t \sim \text{N}(\lambda\zeta_{t-1} + X_t\psi_{t-1}, \omega_t) \tag{3.3}$$

$$\psi_t|\psi_{t-1}, \nu_t \sim \text{N}(\psi_{t-1}, \nu_t). \tag{3.4}$$

Here ζ_t captures the effect of the current and past regression effects, as seen in Equation (3.3). The parameter ψ_t determines the immediate effect X_t has on the quantile. Alternatively the parameter λ is a quantity in the unit interval which represents the memory of the regression effect up to time t. This effect decays at an exponential rate, reducing by a factor of λ at every time step. To see this more explicitly we can derive the transfer function effect k steps ahead, that is:

$$\zeta_{t+k} = \lambda^k\zeta_t + \sum_{r=0}^{k} \lambda^{k-r}\psi_{t+r-1}X_{t+r} + \partial\zeta_{t+k}. \tag{3.5}$$

Thus, the effect of X_t on the quantile at time $t + k$ is $\lambda^k\psi_{t-1}X_t$. This effect is negligible when $\lambda^k|\psi_{t-1}X_t| \leq \epsilon$ for small ϵ. Using this we can derive a series, k_t, representing a lower bound for the number of time steps until the effect of X_t is less than or equal to a fixed ϵ. That is, for $t = 1, \ldots, T$

$$k_t \geq \frac{\log(\epsilon) - \log(|\psi_{t-1}X_t|)}{\log(\lambda)}. \tag{3.6}$$

To complete the model, conjugate priors are available for the additional transfer function parameters; normal conjugate priors for $\zeta_0 \sim N(m_{\zeta_0}, C_{\zeta_0})$ and $\psi_0 \sim$

$N(m_{\psi_0}, C_{\psi_0})$, and a conjugate normal truncated to the unit interval prior for $\lambda \sim N_{(0,1)}(m_\lambda, v_\lambda)$.

3.2.1 MCMC and ISVB Algorithm Augmentations

This transfer function exDQLM can equivalently be rewritten in the form of a standard exDQLM

$$y_t|\gamma, \boldsymbol{\theta}_t, \sigma \sim \text{exAL}_{p_0}(\tilde{\mathbf{F}}_t'\tilde{\boldsymbol{\theta}}_t, \sigma, \gamma) \tag{3.7}$$

$$\tilde{\boldsymbol{\theta}}_t|\tilde{\boldsymbol{\theta}}_{t-1}, \tilde{\mathbf{W}}_t \sim \text{N}(\tilde{\mathbf{G}}_t\tilde{\boldsymbol{\theta}}_{t-1}, \tilde{\mathbf{W}}_t) \tag{3.8}$$

where $\tilde{\mathbf{F}}_t' = (\mathbf{F}_t', 1, 0)$, $\tilde{\boldsymbol{\theta}}_t' = (\boldsymbol{\theta}_t', \zeta_t, \psi_t)$, $\tilde{\mathbf{G}}_t = \text{blockdiag}\left\{\mathbf{G}_t, \begin{pmatrix} \lambda & X_t \\ 0 & 1 \end{pmatrix}\right\}$, and $\tilde{\mathbf{W}}_t = \text{blockdiag}\left\{\mathbf{W}_t, \begin{pmatrix} \omega_t & 0 \\ 0 & \nu_t \end{pmatrix}\right\}$.

Using this representation, the exDQLM MCMC Algorithm 1 can easily be augmented to incorporate the transfer function structure as follows: (1) Replace all $\{\mathbf{F}_t, \boldsymbol{\theta}_t, \mathbf{G}_t, \mathbf{W}_t\}$ with $\{\tilde{\mathbf{F}}_t, \tilde{\boldsymbol{\theta}}_t, \tilde{\mathbf{G}}_t, \tilde{\mathbf{W}}_t\}$, where $\tilde{\mathbf{G}}_t$ will be conditional on $\lambda^{(i)}$; (2) For each iteration i, add an additional step to sample $\lambda^{(i+1)}|\boldsymbol{\zeta}^{(i+1)}, \boldsymbol{\psi}^{(i+1)} \sim N_{(0,1)}(\mu_\lambda, \sigma_\lambda^2)$ with

$$\sigma_\lambda^2 = \left[\sum_{t=1}^{T} \frac{\zeta_{t-1}^{(i+1)2}}{\omega_t} + \frac{1}{v_\lambda}\right]^{-1},$$

$$\mu_\lambda = \sigma_\lambda^2 \left[\sum_{t=1}^{T} \frac{\zeta_{t-1}^{(i+1)}(\zeta_t^{(i+1)} - \psi_t^{(i+1)}X_t)}{\omega_t} + \frac{1}{v_\lambda}\right].$$

Augmenting the ISVB algorithm is not as straight-forward. The random parameter λ within the evolution matrix $\tilde{\mathbf{G}}_t$ compromises our ability to update the state parameter variational distributions within the FFBS while using discount factors to specify $\tilde{\mathbf{W}}_t$. To preserve the ability to utilize discount factors, we propose optimizing the parameter λ with respect to the KL divergence of the one-step-ahead predictive distribution functions as discussed in Sections 2.2.3 and

2.2.4. For optimal λ, say $\tilde{\lambda}$, the ISVB algorithm can be augmented to incorporate the transfer function structure by simply replacing all $\{\mathbf{F}_t, \boldsymbol{\theta}_t, \mathbf{G}_t, \mathbf{W}_t\}$ with $\{\tilde{\mathbf{F}}_t, \tilde{\boldsymbol{\theta}}_t, \tilde{\mathbf{G}}_t, \tilde{\mathbf{W}}_t\}$ where $\tilde{\mathbf{G}}_t = \text{blockdiag}\left\{\mathbf{G}_t, \begin{pmatrix} \tilde{\lambda} & X_t \\ 0 & 1 \end{pmatrix}\right\}$.

3.3 Estimating the 0.85 quantile IVT threshold

3.3.1 Model selection

Consider first a model, M_{01}, in which we estimate the 0.85 quantile with a transfer function component ζ_t to capture the effects of ELI, baseline component $\eta_{1,t}$, and seasonal components $\alpha_{1,t}^{(l)}$, for harmonics $l = 1, \ldots, 6$ of a period of 365 days. More specifically, at each time t we decompose our 0.85 quantile μ_t as follows:

$$\mu_t = \eta_{1,t} + \sum_{l=1}^{6} \alpha_{1,t}^{(l)} + \zeta_t. \tag{3.9}$$

We define the baseline component in the model, $\eta_{1,t}$, with a second order polynomial as follows:

$$\begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_{1,t-1} \\ \eta_{2,t-1} \end{pmatrix} + \boldsymbol{\omega}_t^{\eta}, \quad \boldsymbol{\omega}_t^{\eta} \sim N_2(\mathbf{0}, \mathbf{W}_t^{\eta}) \ . \tag{3.10}$$

This second order polynomial allows the baseline component of the IVT quantile to vary linearly over the extended time frame. Here the system evolution error vectors $\boldsymbol{\omega}_t^{\eta}$, are assumed to be independent over time. We will denote this trend component evolution matrix as $\mathbf{G}^{\eta} = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Further, each harmonic l included in the model is described with a Fourier form representation of cyclical functions,

given as

$$\begin{pmatrix} \alpha_{1,t}^{(l)} \\ \alpha_{2,t}^{(l)} \end{pmatrix} = \begin{pmatrix} \cos(\frac{2\pi}{365}l) & \sin(\frac{2\pi}{365}l) \\ -\sin(\frac{2\pi}{365}l) & \cos(\frac{2\pi}{365}l) \end{pmatrix} \begin{pmatrix} \alpha_{1,t-1}^{(l)} \\ \alpha_{2,t-1}^{(l)} \end{pmatrix} + \boldsymbol{\omega}_t^{\alpha,l}, \quad \boldsymbol{\omega}_t^{\alpha,l} \sim N_2(\mathbf{0}, \mathbf{W}_t^{\alpha,l}) \; . \tag{3.11}$$

We denote the lth seasonal evolution matrix $\mathbf{G}^{\alpha,l} = \begin{pmatrix} \cos(\frac{2\pi}{365}l) & \sin(\frac{2\pi}{365}l) \\ -\sin(\frac{2\pi}{365}l) & \cos(\frac{2\pi}{365}l) \end{pmatrix}$. Again, it is assumed that $\boldsymbol{\omega}_t^{\alpha,l}$ are independent over time, as well as independent of $\boldsymbol{\omega}_t^{\eta}$ for $t = 1, \ldots, T$. For this application we are not interested in higher frequency effects, thus we limit our model to consider only harmonics $l = 1, .., 6$ in the seasonal component.

We choose to model the baseline and seasonal components in this analysis as non-time-varying, thus any variation in the 0.85 quantile from year to year will solely be attributed to the effects of the ELI as input to our transfer function model. This is easily done utilizing component discounting to specify $\mathbf{W}_t$ with discount factor values of 1 (West and Harrison, 2006), which also preserves our ability to update the state parameter using FFBSm. Note that, under this modeling choice, the baseline and seasonal parameters in the state vector are non-time-varying, thus we will omit subscripts t going forward where applicable.

To set the point-mass prior for the scale parameter σ, we considered first the prior suggested in Section 2.7 with point-mass at $\hat{\sigma}_{\gamma=0}^{p_0} \approx 20$. However, we found the ISVB algorithm still struggled to converge so we decrease the location of the point-mass to 15; a decrease of approximately twenty-five percent as suggested in the previous chapter, Section 2.2.5.

The rate parameter λ as well as the discount factors for the evolution of the transfer function parameters, δ_ζ and δ_ψ, are optimized using the KL divergence of the one-step-ahead predictive distribution functions. More specifically, if we denote $[h|\lambda, \delta_\zeta, \delta_\psi]$ as the numerically approximated density of the transformed (by the standard normal CDF) sequence $\{\hat{u}_t\}$ estimated with parameter λ and

discount factors $(\delta_\zeta, \delta_\psi)$, the optimal values will be as follows,

$$\{\tilde{\lambda}, \delta_\zeta, \delta_\psi\} = \underset{\lambda, \delta_\zeta, \delta_\psi}{\arg\min} \int_{-\infty}^{\infty} [h(x)|\lambda, \delta_\zeta, \delta_\psi] \log \frac{[h(x)|\lambda, \delta_\zeta, \delta_\psi]}{\phi(x)} \mathrm{d}x \tag{3.12}$$

with ϕ denoting the standard normal density. Optimal values for M_{01} can be found in Table 3.1.

	M_{01}	M_1
$(\delta_\zeta, \delta_\psi)$	(0.97,0.97)	(0.97,0.97)
$\tilde{\lambda}$	0.37	0.38
γ	-2.39 (-2.40,-2.38)	-2.39 (-2.41,-2.38)
$\eta_{1,1}$	126.93 (125.13,128.64)	127.18 (125.51,128.98)
$\eta_2 \times 10^4$	11.67 (9.61,13.73)	11.58 (9.68,13.57)
pplc	449103.6	447373.1
KL	0.621	0.614

Table 3.1: IVT analysis results for M_{01} and M_1. Optimal δ_ζ, δ_ψ, and $\tilde{\lambda}$. Posterior summaries (format: mean (95% CrI)) for skewness parameter γ, baseline value at $t = 1$, $\eta_{1,1}$, as well as the change in the baseline at each time step, η_2. pplc: Posterior predictive loss criterion under the check loss function. KL: Kullback-Liebler divergences of the one-step-ahead distributions.

Table 3.1 summaries the estimates of the second-order polynomial component with the 95% credible interval (CrI) of the baseline itself at time $t = 1$, $\eta_{1,1}$, as well as the 95% CrI of the change in the baseline at each time step, η_2 (not time-varying). It is clear both the terms are significant. Similarly, Figure 3.2 illustrates 95% CrIs for the seasonal cycles for harmonics $l = 1, \ldots, 6$ for one period of 365 days. Our assessment based on these CrIs indicates the harmonics $1, 2$ and 4 distinct from zero and therefore significant. However, the CrIs of the other 3 components all either overlap with zero or are very close. Keeping in mind variational algorithms underestimate variability, we conclude these harmonics are not significant and omit them in our formal analysis.

Removing the components that are not significant leaves us with a model

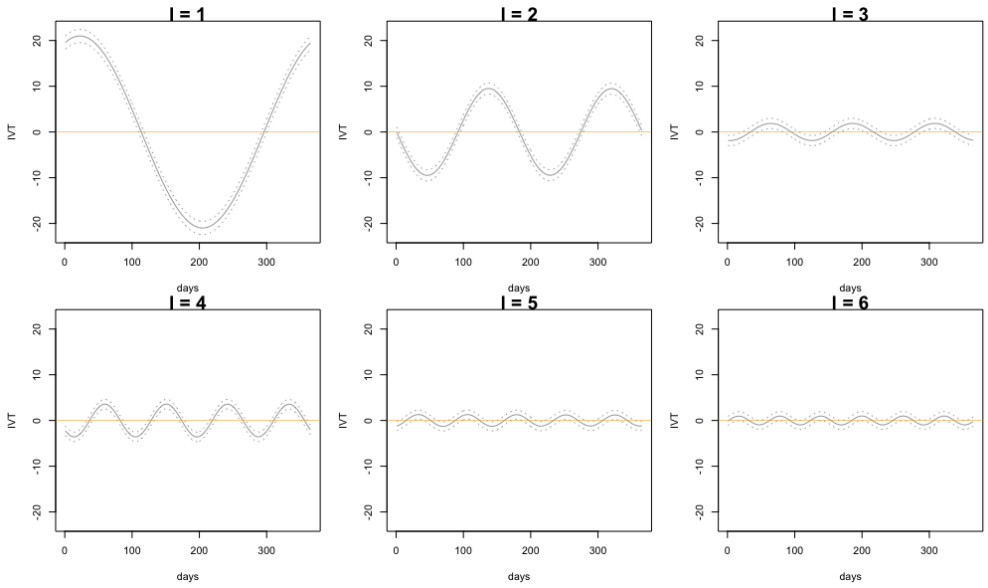

Figure 3.2: M_{01} seasonal components for harmonics $l = 1, 2, \ldots, 6$.

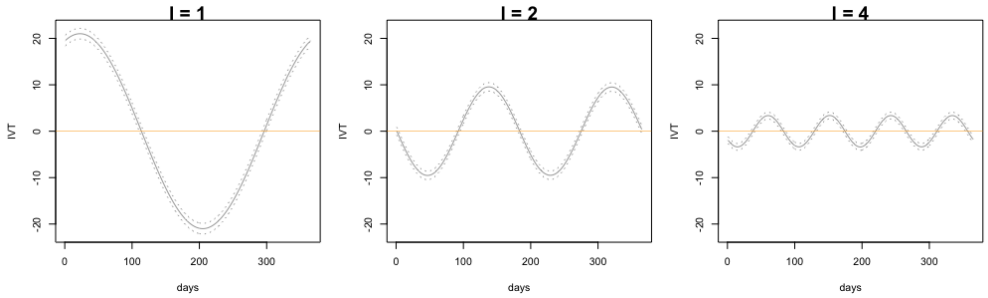

Figure 3.3: M_1 retained seasonal components for harmonics $l = 1, 2, 4$.

comprised of a transfer function component, a second order baseline component and seasonal components for harmonics $l = 1, 2, 4$. Analogous results for this reduced model can be found in Table 3.1 and Figure 3.3 for comparison to M_{01}. It clear the retained components remain unchanged, however the pplc and KL divergence (also seen in Table 3.1) are both lower for this model, suggesting a superior fit to M_{01}. We call this reduced model M_1 and present its results formally in the following section.

3.3.2 IVT analysis

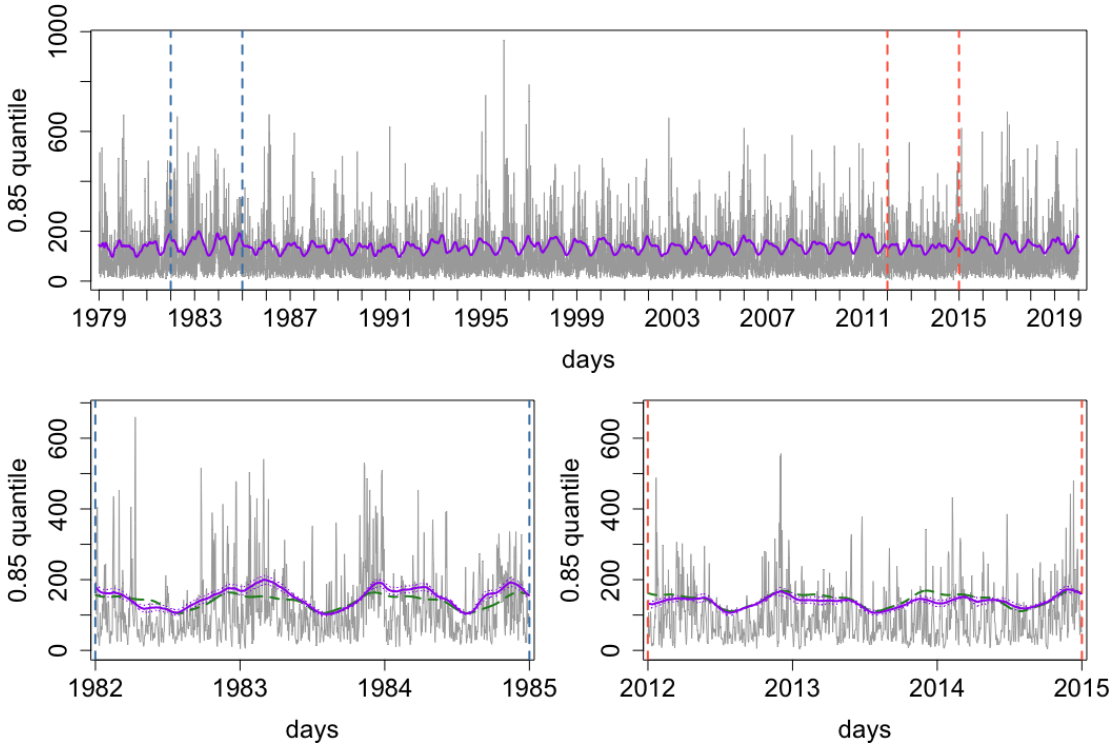

Figure 3.4: Top panel: IVT data with overlaid solid line indicating the M_1 MAP 0.85 quantile threshold of the average daily IVT magnitude in Santa Cruz, CA. The vertical dashed lines enclose the two periods enlarged in the bottom panels. Bottom panels: Years 1982 to 1985 (left, exceptional amount of rain) and years 2012 to 2015 (right, exceptionally dry). The MAP 0.85 quantile from M_1 and M_0 are seen in solid and dashed lines, respectively. The dotted lines surrounding the M_1 MAP estimate indicate the 95% CrIs.

To assess the utility of the transfer funtion component, we fit two separate models to estimate the 0.85 quantile of the IVT magnitude. The first, M_0, is a simplified version of the second, M_1. The simplified M_0 includes only a baseline level and seasonal effects without any input from the ELI time series. Our analysis will focus on the results of M_1, only highlighting features of M_0 for comparison purposes.

As detailed in the previous section, for M_1 we define the structure of the quantile to include a second order baseline, as well as seasonal components for

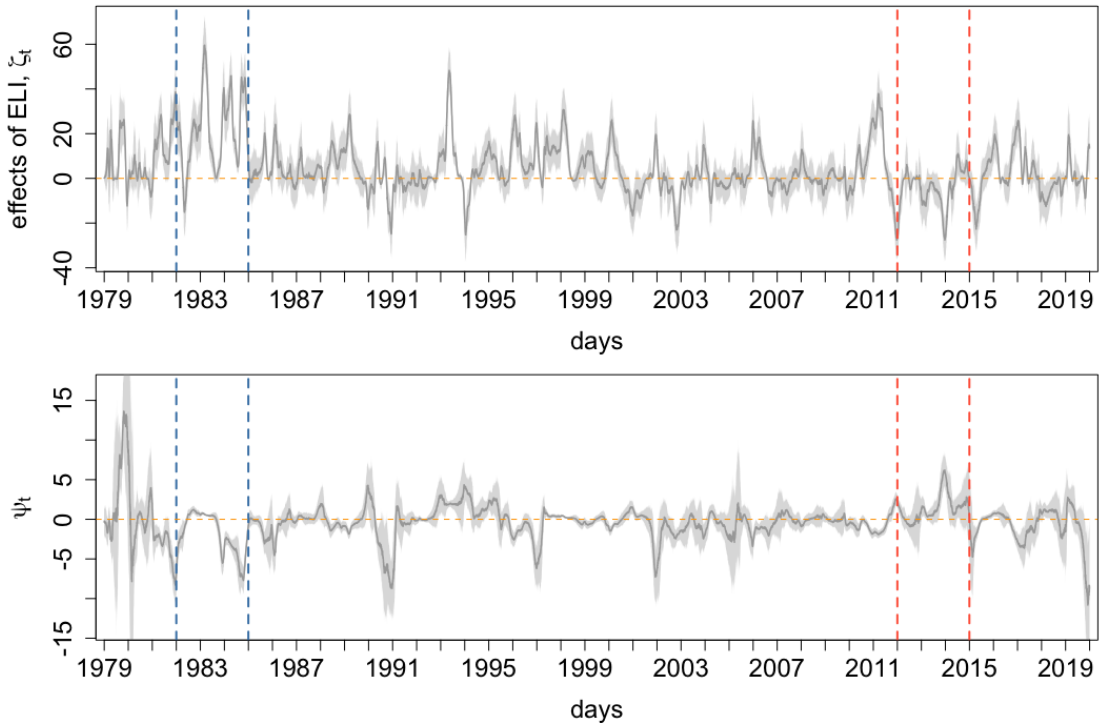

Figure 3.5: Top panel: Effects of ELI captured by the transfer function component, ζ_t. Bottom panel: Instantaneous effects of ELI, ψ_t. In both panels, dark grey lines indicate the MAP estimates. 95% CrI are indicated by the grey shaded regions. Dashed horizontal dashed lines are at zero, for reference. Left vertical dashed lines enclose years 1982 to 1985 in which CA saw an exceptional amount of rain. Right vertical dashed lines enclose 2012 to 2015 in which CA was exceptionally dry.

the annual ($l = 1$), semi-annual ($l = 2$), and quarterly ($l = 4$) harmonics. Using superposition the resulting state vector $\boldsymbol{\theta}_t$, $\mathbf{F}'$, $\mathbf{G}$ and $\mathbf{W}_t$ in Equations (3.1)-(3.4) are defined, respectively, as follows:

$$\boldsymbol{\theta}_t = (\eta_{1,t}, \eta_{2,t}, \alpha^{(1)}_{1,t}, \alpha^{(1)}_{2,t}, \alpha^{(2)}_{1,t}, \alpha^{(2)}_{2,t}, \alpha^{(4)}_{1,t}, \alpha^{(4)}_{2,t})', \tag{3.13}$$

$$\mathbf{F}' = (1, 0, 1, 0, 1, 0, 1, 0), \tag{3.14}$$

$$\mathbf{G} = \text{blockdiag}(\mathbf{G}^{\eta}, \mathbf{G}^{\alpha,1}, \mathbf{G}^{\alpha,2}, \mathbf{G}^{\alpha,4}), \tag{3.15}$$

$$\mathbf{W}_t = \text{blockdiag}(\mathbf{W}^{\eta}_t, \mathbf{W}^{\alpha,1}_t, \mathbf{W}^{\alpha,2}_t, \mathbf{W}^{\alpha,4}_t). \tag{3.16}$$

	M_0	M_1
$(\delta_\zeta, \delta_\psi)$	—	(0.97,0.97)
$\tilde{\lambda}$	—	0.38
γ	-2.47 (-2.48,-2.46)	-2.39 (-2.41,-2.38)
$\eta_{1,1}$	137.82 (137.01,138.83)	127.18 (125.51,128.98)
$\eta_2 \times 10^4$	4.86 (3.61,5.75)	11.58 (9.68,13.57)
A_1	22.97 (22.23,23.63)	20.97 (20.00,22.05)
P_1	-0.45 (-0.48,-0.42)	-0.39 (-0.45,-0.34)
A_2	9.67 (8.93,10.38)	9.58 (8.66,10.47)
P_2	1.53 (1.46,1.57)	1.52 (1.44,1.57)
A_4	3.54 (2.92,4.17)	3.41 (2.63,4.15)
P_4	-1.02 (-1.19,-0.80)	-1.02 (-1.25,-0.81)
pplc	450214.7	447373.1
KL	0.856	0.614
run-time	6.87	7.92

Table 3.2: IVT analysis results for M_0 and M_1. Optimal δ_ζ, δ_ψ, and $\tilde{\lambda}$, as discussed in Section 3.3.1. Posterior summaries (format: mean (95% CrI)) for skewness parameter γ, baseline at $t = 1$ $\eta_{1,1}$, the change in the baseline at each time step η_2, annual amplitude A_1, annual phase P_1, semi-annual amplitude A_2, semi-annual phase P_2, quarterly amplitude A_4, quarterly phase P_4. pplc: Posterior predictive loss criterion under the check loss function. KL: Kullback-Liebler divergences of the one-step-ahead distributions. Run-time: ISVB run-times in minutes.

In addition to the baseline and seasonal components, in M_1 we utilize the exponentially decaying transfer function exDQLM as specified in Equations (3.1)-(3.4) to capture both the immediate and lagged effects of ELI on the 0.85 quantile. We complete the models with conjugate priors, where applicable; $\boldsymbol{\theta}_0 \sim \text{N}(\mathbf{m}_0, \mathbf{C}_0)$ with $\mathbf{m}_0 = \mathbf{0}$ and $\mathbf{C}_0 = 100I_q$, $\zeta_0 \sim N(m_{\zeta_0}, C_{\zeta_0})$ and $\psi_0 \sim N(m_{\psi_0}, C_{\psi_0})$ with $m_{\zeta_0} = m_{\psi_0} = 0$ and $C_{\zeta_0} = C_{\psi_0} = 10$, and $\gamma \sim t_{(-5.137,0.213)}(0,1)$ with 1 degree of freedom. The parameter σ is fixed at a value of 15 to facilitate convergence, as discussed in the previous section. Lastly, in M_1, the parameter λ as well as the discount factors for the evolution of ζ_t and ψ_t, δ_ζ and δ_ψ respectively, are optimized using the KL divergence of the one-step-ahead predictive distribution functions as discussed in Section 3.3.1. Optimal λ, δ_ζ, δ_ψ can be found in Table

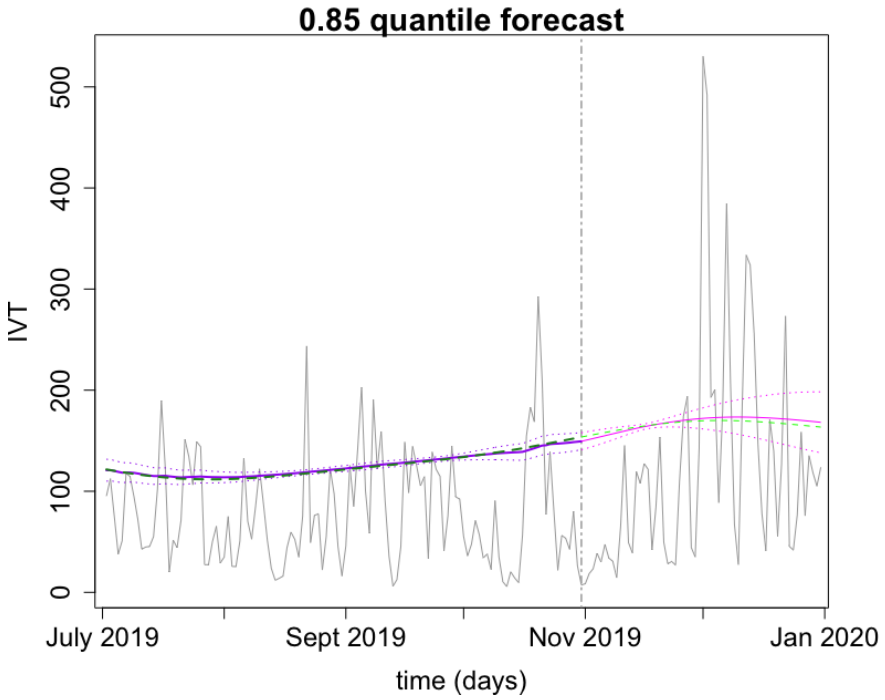

Figure 3.6: 60-step-ahead quantile forecast beginning November 2, 2019 through December 31, 2019 overlaid on the IVT magnitude data. The vertical dot-dashed line is at the beginning of the forecast period, November 2, 2019, for reference. From M_1: Solid lines indicate mean estimates and dotted lines 95% CrI, with the filtered estimates seen leading up to November 2$^{\text{nd}}$ and the forecast estimates beyond November 2$^{\text{nd}}$. From M_0: The dashed lines indicate the mean estimates, again with the filtered and forecast estimates seen before and after November 2$^{\text{nd}}$, respectively.

3.2.

We apply our ISVB algorithm to estimate the 0.85 quantile of the daily IVT magnitude in Santa Cruz, CA from 1979 to 2019. The ISVB algorithm was implemented in the R programming language (R Core Team, 2013) on a personal computer with a 2.5 GHz Intel Core i5 processor. For this time series of length 14965 the ISVB computation times are under eight minutes for both models; exact times can be found in Table 3.2. Also seen in Table 3.2, the CrIs of skewness parameter γ are distinct from 0 for both M_0 and M_1, thus justifying the added

flexibility of our exDQLM versus the DQLM in this application. A full comparison of M_1 with an analogous transfer function DQLM can be found in the following section.

Figure 3.4 illustrates the MAP 0.85 quantile of M_1 for the entire time period, as well as the MAP estimates from both models for the two time periods, 1982 through 1985 and 2012 through 2015, in which CA saw drastically different amounts of precipitation. It can be seen that the CrI of M_1 are often distinct from M_0, with the M_1 quantile generally higher than the M_0 quantile when CA experienced an exceptional amount of rain and the opposite true when CA was exceptionally dry.

From the posterior estimates of the annual and semi-annual harmonic components of the models we compute the amplitude and phase, $A_l = \sqrt{(\alpha_1^l)^2 + (\alpha_2^l)^2}$ and $P_l = \arctan(-\alpha_2^l/\alpha_1^l)$ respectively. Posterior summaries of these as well as the baseline parameters can be found in Table 3.2. It can be seen that the value of the second-order polynomial trend component at $t = 1$, $\eta_{1,1}$, is significantly larger in M_0 than M_1. Conversely, the parameter which describes the rate of change per time step (non-time-varying), η_2, is significantly smaller in M_0 than M_1. This suggests the inclusion ELI in M_1 accounts for some of the long-term variability seen in the 0.85 quantile from 1979 through 2019. Differences can also be seen in the annual components where the amplitude of M_0 is significantly larger than M_1 and the phase of M_0 is significantly smaller than that of M_1. Alternatively, the semi-annual and quarterly harmonic components are indistinct between the two models. This suggests there is a substantial amount of variability in the 0.85 quantile that can be associated with the ELI time series specifically on the annual time scale, with the distinction less clear at the semi-annual and quarterly scales.

The amount of variability in the 0.85 quantile attributed to the effects of

ELI captured with transfer function component in M_1, ζ_t, are seen in Figure 3.5. The effects of ELI are overall significant and are dramatically more pronounced between 1982 and 1985 than between 2012 and 2015. In particular, a majority of the effects between 1982 and 1985 (in which CA received heavy precipitation) are distinctly positive whereas the effects between 2012 to 2015 (when drought was severe) are negative or not significant. The instantaneous effects of ELI at time t, ψ_t, also exhibit very different behavior in the two time periods, seen in Figure 3.5. Upon computing the series k_t from Equation (3.6) for $\epsilon = 1\text{e}{-3}$ (not pictured), we find the lagged effects of ELI persist for around 8.5 days, on average.

To assess the predictive value added by the transfer function component capturing the effects of ELI in M_1, we compare the pplc and KL divergence of the one-step-ahead forecast distributions for M_1 to those of M_0, seen in Table 3.2. M_1 is favored with smaller values of both model comparison criterion, suggesting the inclusion of ELI improves both the predictive and forecasting power of M_1. As further model validation, a thorough examination of the correlation and distributional shape of the one-step-ahead predictive distributions (discussed in the previous chapter Section 2.2.3) is presented Section 3.3.4. To further investigate the information added by ELI, we can examine the k-step-ahead quantile forecast distributions. More specifically, for each time t we can compute the k-step-ahead future distribution of state vector

$$\tilde{\boldsymbol{\theta}}_{t+k}|\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}, y_1, \ldots, y_t \sim \text{N}(\mathbf{a}_t(k), \mathbf{R}_t(k)) \tag{3.17}$$

where $\mathbf{a}_t(k) = \tilde{\mathbf{G}}_{t+k}\mathbf{a}_t(k-1)$, $\mathbf{R}_t(k) = \tilde{\mathbf{G}}_{t+k}\mathbf{R}_t(k-1)\tilde{\mathbf{G}}'_{t+k} + \tilde{\mathbf{W}}_{t+k}$, $\mathbf{a}_t(0) = \mathbf{m}_t$, and $\mathbf{R}_t(0) = \mathbf{C}_t$, with $\mathbf{m}_t$ and $\mathbf{C}_t$ denoting the filtered mean and covariance of $\tilde{\boldsymbol{\theta}}_t$, respectively. $\tilde{\mathbf{W}}_{t+k}$ can again be specified using discount factors such that the i^{th} component is defined $\tilde{\mathbf{W}}_{i,t+k} = \frac{(1-\delta_i)}{\delta_i}\tilde{\mathbf{G}}_{i,t+k}\mathbf{R}_{i,t}(k-1)\tilde{\mathbf{G}}'_{i,t+k}$ where $\tilde{\mathbf{G}}_{i,t+k}$

and $\mathbf{R}_{i,t}(k-1)$ denote the i^{th} components of $\tilde{\mathbf{G}}_{t+k}$ and $\mathbf{R}_t(k-1)$, respectively, and δ_i denotes i^{th} discount factor. From this k-step-ahead future distribution of state vector, we can derive the k-step-ahead future marginal distribution of the quantile,

$$\tilde{\mathbf{F}}'_{t+k}\tilde{\boldsymbol{\theta}}_{t+k}|\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}, y_1, \ldots, y_t \sim \mathrm{N}(\tilde{\mathbf{F}}'_{t+k}\mathbf{a}_t(k), \tilde{\mathbf{F}}'_{t+k}\mathbf{R}_t(k)\tilde{\mathbf{F}}_{t+k}). \tag{3.18}$$

The posterior means and 95% CrIs of the 60-steps-ahead quantile distributions can be seen in Figure 3.6. The MAP quantile forecast of M_1, which takes the ELI into consideration, suggests the 0.85 quantile will be very similar to the seasonal average projected by M_0 in this particular time period.

3.3.3 Comparison with the DQLM special case

Next, we illustrate the differences between 0.85 quantile IVT threshold estimated using the DQLM versus those from the exDQLM, e.g. M_1. The results from the DQLM we will call M_2 going forward.

M_1 and M_2 are estimated using the transfer function exDQLM with identical components; a second-order polynomial baseline and seasonal effects for harmonics $l = 1, 2, 4$. The rate parameter λ used in M_2 is optimized using the KL divergence of the one-step-ahead predictive distribution functions as discussed in Section 3.3.1, the value of which can be found in Table 3.3. However, the discount factors of M_2 are fixed to the optimal values from M_1 to facilitate comparison.

Figure 3.7 illustrates the MAP 0.85 quantile of the models for the entire time period from 1979 through 2019, as well as the 95% CrI for the two time periods, 1982 through 1985 and 2012 through 2015, in which CA saw drastically different amounts of precipitation. It can be seen that the CrI of M_1 are distinct from M_2 with the M_2 quantile generally higher than the M_1 quantile, particularly in the

	M_1	M_2
$(\delta_\zeta, \delta_\psi)$	(0.97,0.97)	(0.97,0.97)
$\tilde{\lambda}$	0.38	0.42
σ	15 (fixed)	20.13 (19.84,20.42)
γ	-2.39 (-2.41,-2.38)	0 (fixed)
$\eta_{1,1}$	127.18 (125.51,128.98)	132.85 (130.85,134.67)
$\eta_2 \times 10^4$	11.58 (9.68,13.57)	25.46 (23.32,28.07)
A_1	20.97 (20.00,22.05)	41.44 (39.90,42.89)
P_1	-0.39 (-0.45,-0.34)	-0.22 (-0.26,-0.18)
A_2	9.58 (8.66,10.47)	15.71 (14.60,17.04)
P_2	1.52 (1.44,1.57)	1.22 (1.16,1.30)
A_4	3.41 (2.63,4.15)	4.76 (3.76,5.79)
P_4	-1.02 (-1.25,-0.81)	-0.72 (-0.90,-0.47)
pplc	447373.1	1119638
KL	0.614	0.958
run-time	7.92	4.23

Table 3.3: IVT analysis results for M_1 and M_2. Identical values of (δ_ζ, δ_ψ) used in both models and optimized within M_1. Rate parameter $\tilde{\lambda}$ optimized within each model. The scale parameter σ is fixed for M_1 as discussed in Section 2.7, and the skewness parameter γ is fixed to 0 for M_2 (reducing the model to the DQLM). Posterior summaries (format: mean (95% CrI)) for M_2 scale parameter σ, M_1 skewness parameter γ, baseline at $t = 1$ $\eta_{1,1}$, the change in the baseline at each time step η_2, annual amplitude A_1, annual phase P_1, semi-annual amplitude A_2, semi-annual phase P_2, quarterly amplitude A_4, quarterly phase P_4. pplc: Posterior predictive loss criterion under the check loss function. KL: Kullback-Liebler divergences of the one-step-ahead distributions. Run-time: ISVB run-times in minutes.

winter months. Given what was seen in the simulation study of Section 2.3 of the previous chaper, this result is not surprising. The DQLM is substantially affected by the extreme observations of IVT which occur most commonly in the winter months.

Differences between the two models can also be seen in Table 3.3. Every component of the model is affected when using the DQLM. Most noticeably, these include the initial baseline value $\eta_{1,1}$ and the baseline increase over time η_2, as well as all seasonal amplitudes. The phases of the seasonal components however

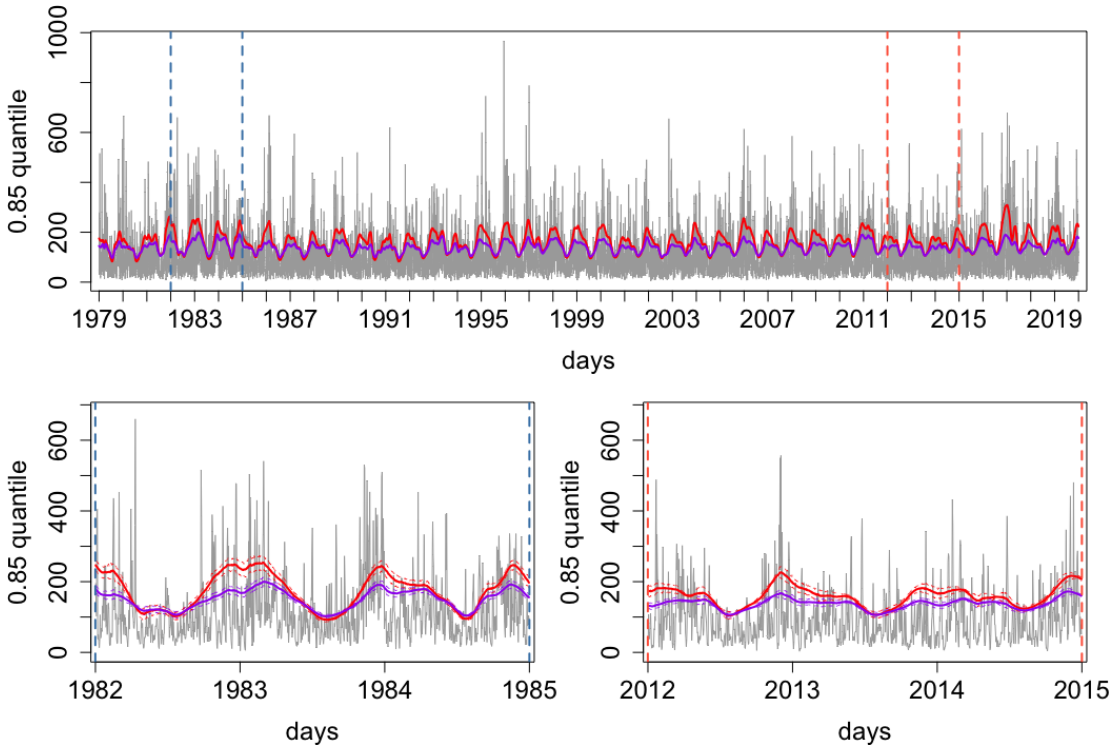

Figure 3.7: Top panel: In purple, M_1 MAP 0.85 quantile threshold of the average daily IVT magnitude in Santa Cruz, CA. In red, M_2 MAP 0.85 quantile. The blue vertical lines enclose a period in which CA saw an exceptional amount of rain. The red vertical lines enclose a period in which CA was exceptionally dry. Bottom panels: Years 1982 to 1985 (left, exceptional amount of rain) and years 2012 to 2015 (right, exceptionally dry). The MAP 0.85 quantile from M_1 and M_2 are seen in purple and red, respectively, with 95% CrIs indicated with dashed lines.

are barely distinct for the annual and semi-annual components, and the 95% CrI of the quarterly phases overlap.

The amount of variability in the 0.85 quantiles attributed to the effects of ELI captured with transfer function components, ζ_t, are seen in Figure 3.8 for both M_1 and M_2. The instantaneous effects of ELI at time t, ψ_t, are also seen in Figure 3.8. The behavior of both parameters at each time point is very similar between the two models. However, as expected the effects of ELI are exaggerated by the DQLM in M_2.

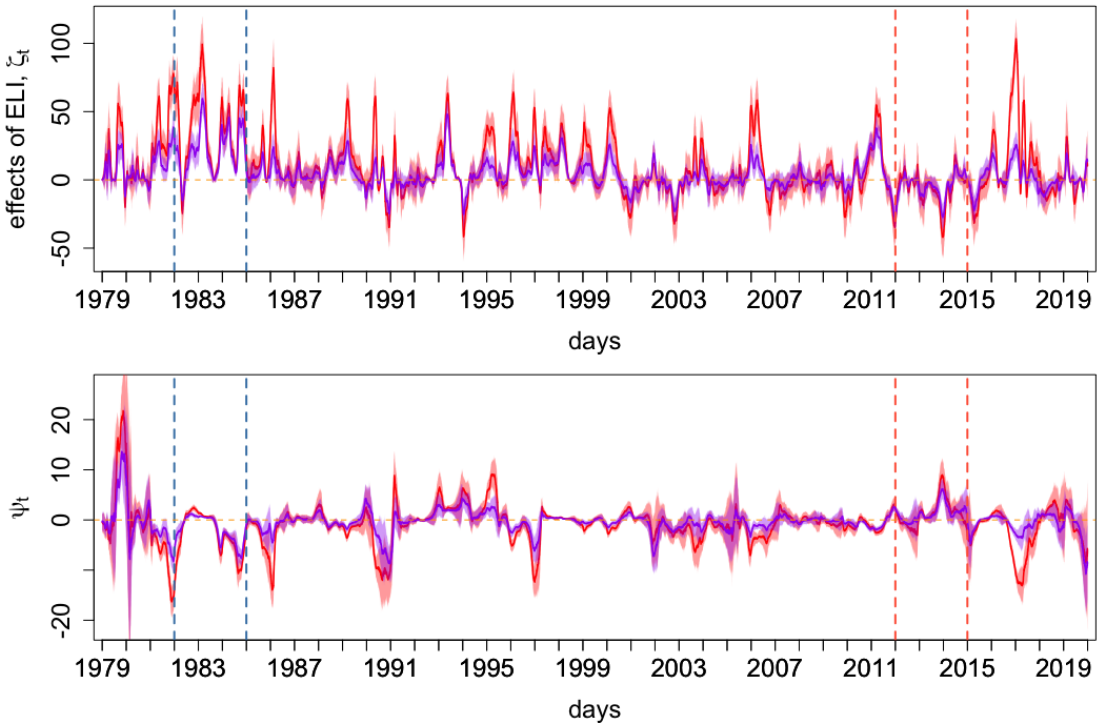

Figure 3.8: Top panel: Effects of ELI captured by the transfer function component, ζ_t. Bottom panel: Instantaneous effects of ELI, ψ_t. In both panels, M_1 is seen in purple and M_2 is seen in red. Solid lines indicate the MAP estimates. 95% CrI are indicated by the shaded regions. Dashed horizontal orange lines are at zero, for reference. Blue vertical lines enclose years 1982 to 1985 in which CA saw an exceptional amount of rain. Red vertical lines enclose 2012 to 2015 in which CA was exceptionally dry.

To further investigate the differences between the exDQLM and DQLM, we can examine k-step-ahead quantile forecast distributions in Equation (3.18). The posterior means and 95% CrIs of these distributions for 60-steps-ahead can be seen in Figure 3.9. The MAP quantile forecast of M_2, suggests the 0.85 quantile will be dramatically higher than the 0.85 quantile projected by M_1.

The predictive value of the models, as well as the overall fit, can be compared using the pplc and KL divergence of the one-step-ahead forecast distributions. These values for M_1 and M_2 can be found in Table 3.3. M_1 is favored with

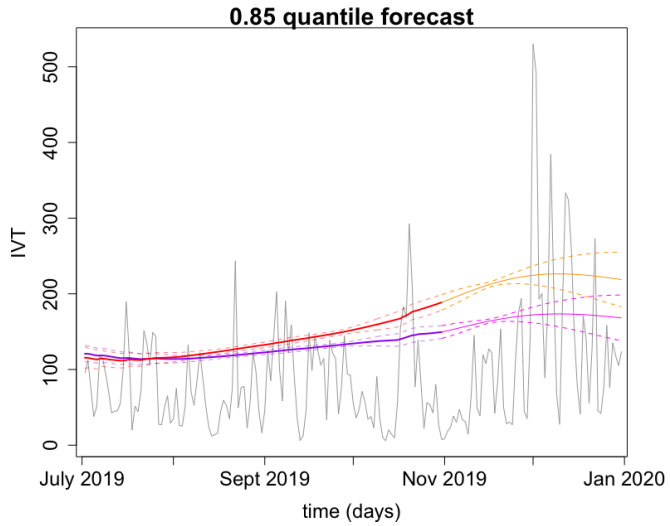

Figure 3.9: 60-step-ahead quantile forecast beginning November 2, 2019 through December 31, 2019. From M_1: The solid magenta line indicates the forecast means and the dashed magenta indicate the 95% CrI of the forecasted 0.85 quantile. Also included are the filtered means and 95% CrI up to November 2, 2019 in purple solid and dashed lines, respectively. From M_2: The solid orange line indicates the forecast means and the dashed orange lines the 95% CrIs. The filtered means and 95% CrI up to November 2, 2019 in red solid and dashed lines, respectively. The IVT magnitude data is seen in grey.

smaller values of both model comparison criterion. Our exDQLM improves on both the predictive and forecasting power of the DQLM, and provides a method for quantifying the relationship between IVT and ELI at the 0.85 quanitle that is robust with respect to extreme observations.

3.3.4 Model Validation

In addition to the KL divergence and pplc in Table 3.2, we can further assess models M_0, M_1, and now M_2 by taking a closer look at the one-step-ahead predictive distributions $u_t = \Phi(y_t|D_{t-1}, \mathbf{v}, \mathbf{s}, \sigma, \gamma)$ introduced in Section 2.2.3.

We diagnose the model performance through the distributional shape of the estimated sequence $\{\hat{u}_t\}$. Transforming the values with a standard normal inverse

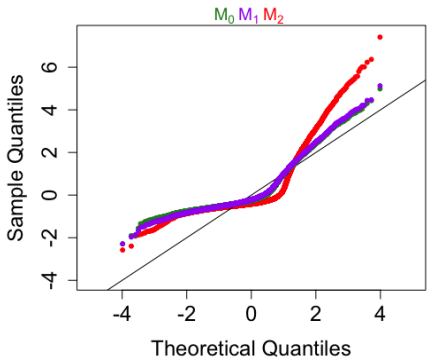

Figure 3.10: QQ-plots of the transformed one-step-ahead predictive distribution functions. M_0 is seen in green, M_1 in purple, and M_2 in red.

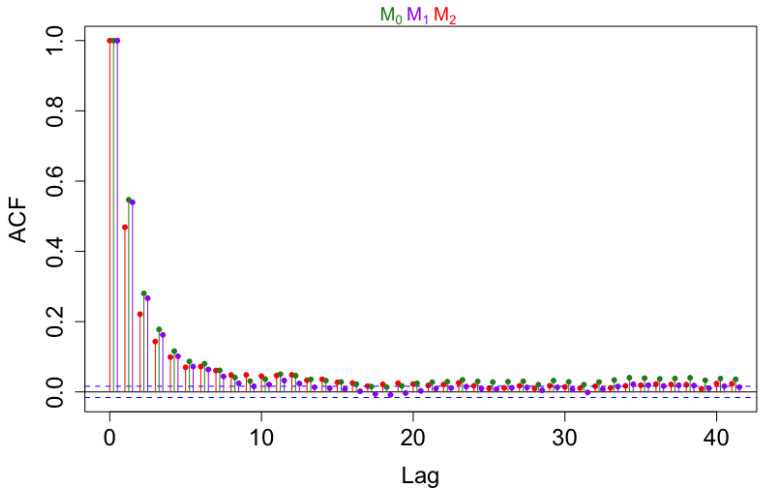

Figure 3.11: ACF plots of the one-step-ahead predictive distribution functions. M_0 is seen in green, M_1 in purple, and M_2 in red.

CDF allows for examination of the distribution shape with a normal QQ-plot, results of which can be found in Figure 3.10. M_0 and M_1, the exDQLMs, produce very similar results to the point that they are difficult to visually distinguish. It is clear, even with the added flexibility of the exDQLM, we are not able to capture the entire structure of the data. While the exAL improves upon the AL which is well known to suffer from model misspecification (Komunjer, 2005), a certain level of misspecification is unavoidable. We are however able to improve upon the

results of the DQLM, which deviates dramatically from the theoretical quantile of the standard normal, particularly in the upper tail of the distribution. The divergence of these transformed values from a standard normal is summarized by the KL divergence in Table 3.3.

Next, we diagnose the model performances through the correlation of the estimated sequence $\{\hat{u}_t\}$. This is done using the ACF plots seen in Figure 3.11. It can be seen that the sequence from the exDQLM without a transfer function component (M_0) has the most persistent auto-correlation. Alternatively, the sequences of both transfer function models (M_1 and M_2) have auto-correlations which have decayed by lag of approximately 20. Although all models exhibit some left over temporal dependence, the correlation of the sequence estimated by the transfer function exDQLM (M_1) decays the fastest with a auto-correlation of approximately zero by a lag of 15.

Finally, we can also use the standardized forecast errors,

$$\frac{y_t - \mathbb{E}[y_t|D_{t-1}, \mathbf{v}, \mathbf{s}, \sigma, \gamma]}{\sqrt{\text{Var}(y_t|D_{t-1}, \mathbf{v}, \mathbf{s}, \sigma, \gamma)}},$$

seen in Figure 3.12 as an additional graphical aid in examining distribution at specific time points. The distributions of these forecast are not available in closed form, however we can observe the scale of the residuals is the largest for M_2, the DQLM. Further, we can observe M_0, which does not include the effects of ELI, has an apparent seasonal pattern in its lower tail which is not present in the other two models. Although none of the models are able to fully capture the entire structure of the IVT time series, it is clear the transfer function exDQLM (M_1) improves upon the transfer function DQLM (M_2) and the inclusion of ELI accounts for a portion of the structure omitted in M_0. Inclusion of more components or covariates may improve this further, however we would like to emphasize the

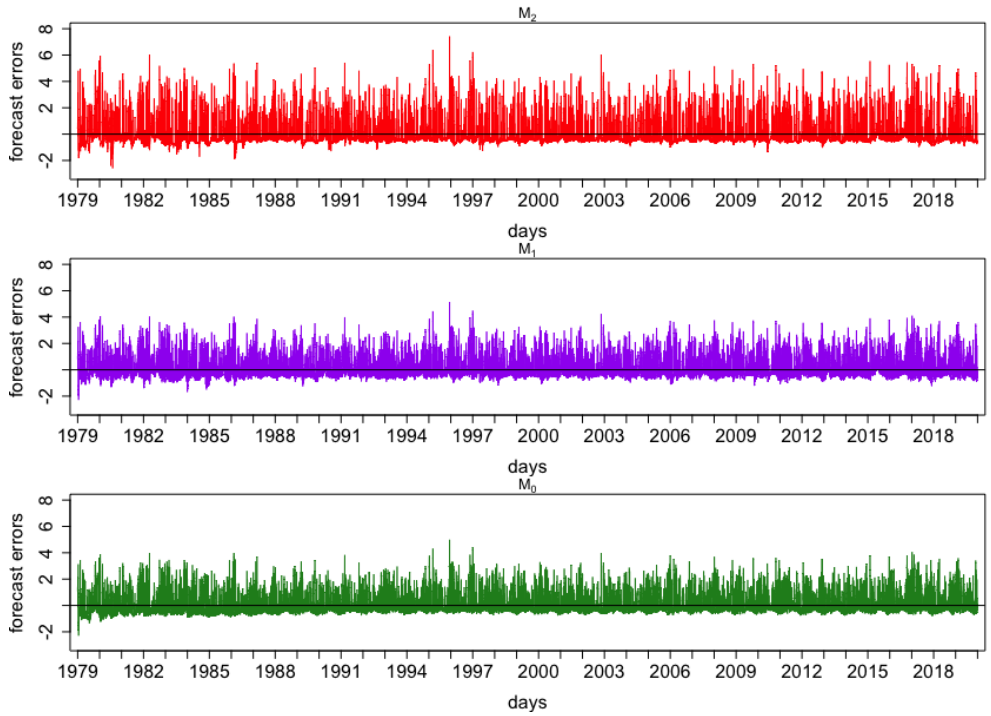

Figure 3.12: MAP estimates of the standard one-step-ahead predictive forecast errors. M_0 is seen in green, M_1 in purple, and M_2 in red.

exDQLM provides a superior model fit to the DQLM regardless of the structure of the quantile.

3.4 Conclusion

In contrast to current schemes, our transfer function exDQLM develops a straight-forward method for quantifying non-linear relationships between a response and input at a specified quantile. We illustrated the utility of our methods in the analysis of the ERA5 IVT magnitude 0.85 quantile threshold in Santa Cruz, CA; an analysis made possible by our ISVB algorithm. In contrast to many current thresholding approaches, estimating the 0.85 quantile with our model provides rich inference about the structure of the time series and thus enhances the tools

for characterization of ARs. The results demonstrated not only significant long-term variability and seasonality, but also a significant non-linear relationship with the climate index ELI captured by the transfer function component of our model. Through several model checking criteria, we were able to show the inclusion of ELI in the model was advantageous both in forecast and predictive accuracy. We saw the effects of ELI on the 0.85 quantile varied substantially between two time periods which experienced drastically different amounts of precipitation; results that are relevant to understanding the roll of IVT magnitude and ARs in the global water cycle and regional weather.

Chapter 4

exdqlm: An R Package for Estimation and Analysis of Flexible Dynamic Quantile Linear Models

In this chapter, we present the (R Core Team, 2013) package **exdqlm**. We discuss in detail the functionalities of the package and illustrate implementation of the routines with step-by-step analyses of several real datasets.

4.1 Introduction

The purpose of the R (R Core Team, 2013) package **exdqlm** is to provide an accessible and comprehensive tool for flexible Bayesian estimation of a single dynamic quantile. Implementing the methodology presented in Chapters 2 and 3, the package includes routines for model building, estimation and diagnostics. Both the MCMC and ISVB algorithms are implemented within the package, as well as the transfer function extension to the exDQLM. The package is protected under the MIT License, and is available from the Comprehensive R Archive Network

(CRAN) at https://CRAN.R-project.org/package=exdqlm.

In the remainder of this chapter we introduce the package as follows. In Section 4.2, we briefly summarize the modeling framework, including the algorithms used for efficient posterior inference and methods for including nonlinear relationships. In Section 4.3 we detail the model diagnostics implemented in the package. For readability some details presented in Chapters 2 and 3 are also included in these sections. Then, in Section 4.4, we walk through three step-by-step examples on how to use the package. Finally, in Section 4.5, we present a summary of the package capabilities.

4.2 Extended dynamic quantile linear models

The package **exdqlm** implements the dynamic quantile models defined in Chapter 2 for inference on a single p_0-th quantile. That is, for a set of scalar time-evolving responses, y_t, and $t = 1, \ldots, T$, an exDQLM is defined by:

$$\text{Observation equation:} \quad y_t = \mathbf{F}_t'\boldsymbol{\theta}_t + \epsilon_t, \qquad \epsilon_t \sim \text{exAL}_{p_0}(0, \sigma, \gamma) \tag{4.1}$$

$$\text{System equation:} \quad \boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \qquad \boldsymbol{\omega}_t \sim \text{N}(0, \mathbf{W}_t). \tag{4.2}$$

As presented in Section 2.2, the mixture representation of the exAL can be exploited to rewrite the exDQLM as the following hierarchical model for $t = 1, \ldots, T$:

$$y_t|\boldsymbol{\theta}_t, \sigma, \gamma, v_t, s_t \sim \text{N}(y_t|\mathbf{F}_t'\boldsymbol{\theta}_t + C(p,\gamma)\sigma|\gamma|s_t + A(p)v_t, \sigma B(p)v_t) \tag{4.3}$$

$$v_t, s_t|\sigma \sim \text{Exp}(v_t|\sigma)\text{N}^+(s_t|0,1) \tag{4.4}$$

$$\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \mathbf{W}_t \sim \text{N}(\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t). \tag{4.5}$$

Input list	`model`				`PriorSigma`		`PriorGamma`		
Symbol	$\mathbf{F}_{1:T}$	$\mathbf{G}_{1:T}$	$\mathbf{m}_0$	$\mathbf{C}_0$	a_σ	b_σ	m_γ	s_γ	ν_γ
List component	`FF`	`GG`	`m0`	`C0`	`a_sig`	`b_sig`	`m_gam`	`s_gam`	`df_gam`
Default	-	-	-	-	2.1	1.1	0	1	1

Table 4.1: Translation from mathematical symbols to parameters used **exdqlm** functions, as well as default values where applicable.

This hierarchical form is leveraged for the posterior inference implemented within the package. Table 4.1 shows the inputs that define the structure of the dynamic model; $\mathbf{F}_t$ and $\mathbf{G}_t$ for $t = 1, \ldots, T$.

4.2.1 Prior specification

Under a Bayesian formulation and given p_0, we specify priors for the initial state vector $\boldsymbol{\theta}_0$, scale parameter σ and skewness parameter γ. For $\boldsymbol{\theta}_0$, we assume a q-variate conjugate normal prior $\boldsymbol{\theta}_0 \sim \mathrm{N}(\mathbf{m}_0, \mathbf{C}_0)$. For σ, we assume a conjugate inverse-gamma prior denoted as $\mathrm{IG}(a_\sigma, b_\sigma)$. In many cases a strong prior on σ is necessary to guarantee fast posterior convergence and estimation, which will be illustrated in the examples of Section 4.4. For the prior on γ, we place a Student-t distribution truncated to the interval (L, U) as defined in Section 2.1, i.e. $\gamma \sim \mathrm{t}_{(L,U)}(m_\gamma, s_\gamma)$ with ν_γ degrees of freedom. The prior parameters used in the **exdqlm** functions, their corresponding mathematical symbols, and default values are shown in Table 4.1.

4.2.2 Posterior estimation

The hierarchical representation of the exDQLM facilitates posterior simulation using the MCMC algorithm, detailed in Section 2.2.1. This is implemented in the function `exdqlmMCMC`. Gibbs sampling steps are used to update the sets of latent variables s_t and v_t for $t = 1, \ldots, T$. The dynamic regression coefficients

are sampled using a forward filtering backwards sampling algorithm (Carter and Kohn, 1994; Frühwirth-Schnatter, 1994). Finally, parameters γ and σ are updated jointly with a random-walk Metropolis-Hastings (MH) step on the rescaled logit- and log-scale, respectively. The covariance matrix used to sample the parameters (Σ_{MH}) can be specified by the user and/or automatically computed from the burned sample. An example of this functionality can be found in Section 4.4.1. We perform a total of $N_{BURN} + N_{MCMC}$ MCMC iterations, discarding the first N_{BURN} resulting in N_{MCMC} posterior samples. The output of the `exdqlmMCMC` function includes the posterior samples of all parameters $(\boldsymbol{\theta}_{1:T}, v_{1:T}, s_{1:T}, \gamma, \sigma)$, a list containing the parameters of the filtered and smoothed distributions of the state vector $\boldsymbol{\theta}_{1:T}$, the covariance matrix Σ_{MH} and the acceptance rate of the MH step.

We include in the package the ISVB routine detailed in Section 2.2.2, implemented in `exdqlmISVB`, for fast and accurate approximate posterior estimation. For the exDQLM, the variational distributions are recognizable closed-form distributions with the exception of the joint variational distribution of σ and γ, which is approximated at each iteration using importance sampling (IS) with N_{IS} particles. The parameters of the remaining distributions are computed directly at each iteration, with the dynamic regression coefficients updated using a forward filtering backwards smoothing algorithm. After the algorithm reaches convergence with tolerance of ϵ_{tol}, N_{SAMP} samples are drawn from the variational distributions, resulting in an approximate sample of the posterior distribution. The output of the `exdqlmISVB` function includes the approximate posterior samples as well as lists containing the parameters of the variational distributions for all parameters $(\boldsymbol{\theta}_{1:T}, v_{1:T}, s_{1:T}, \gamma, \sigma)$. For γ and σ the list also contains the IS weights and samples.

Table 4.2 shows the input parameters used in the functions discussed in this

exdqlm function	`exdqlmMCMC`			`exdqlmISVB`		
Symbol	Σ_{MH}	N_{BURN}	N_{MCMC}	ϵ_{tol}	N_{IS}	N_{SAMP}
Input	`Sig.mh`	`n.burn`	`n.mcmc`	`tol`	`n.IS`	`n.samp`
Default	$0.05\mathbf{I}_2$	2000	1500	0.1	500	200

Table 4.2: Translation from mathematical symbols to parameters used in functions `exdqlmMCMC` and `exdqlmISVB`, as well as default values. Here $\mathbf{I}_2$ denotes the identity matrix of dimension 2.

section, their mathematical symbols and default values.

4.2.3 Specification of the evolution covariance via discount factors

Both estimation algorithms, `exdqlmMCMC` and `exdqlmISVB`, utilize discount factors to specify the time-evolving covariance matrices $\mathbf{W}_t$ found in Equation (4.2). The discount factor δ can be specified by the user via function parameter `df`. Selection of discount factors is typically done by optimizing some model checking criterion. We suggest using the Kullback-Leibler (KL) divergence of the one-step-ahead predictive distribution functions presented in Section 2.2.3. A brief example of this method for discount factor selection can be found in Section 4.4.2.

Discounting by component of the state-space model is also a functionality of the estimation algorithms. Suppose the state vector is comprised of h components $\boldsymbol{\theta}_{it}$ of dimension q_i for $i = 1, \ldots, h$ such that $\boldsymbol{\theta}'_t = (\boldsymbol{\theta}'_{1t}, \ldots, \boldsymbol{\theta}_{ht})'$ and $\sum_{i=1}^{h} q_i = q$. If $\mathbf{W}_{it}$ denotes the i^{th} component of the evolution covariance matrix such that $\mathbf{W}_t = \text{blockdiag}\{\mathbf{W}_{1t}, \ldots, \mathbf{W}_{ht}\}$ and δ_i denotes a discount factor for the i^{th} component, we can define the evolution variance matrix by component as $\mathbf{W}_{it} = \frac{1-\delta_i}{\delta_i}\mathbf{G}_{it}\mathbf{C}_{i,t-1}\mathbf{G}'_{it}$. Here $\mathbf{G}_{it}$ and $\mathbf{C}_{i,t-1}$ denote the i^{th} components of $\mathbf{G}_t$ and $\mathbf{C}_{t-1}$, respectively. The function parameters `df` and `dim.df` allow users to specify a set of component discount factors $(\delta_1, \ldots, \delta_h)$ and their dimensions $(q_1, \ldots, q_h)$, respectively, within the state-space model. Examples of this feature can be found

Symbol	$\tilde{t}$	k	$\mathbf{F}_{(\tilde{t}+1):(\tilde{t}+k)}$	$\mathbf{G}_{(\tilde{t}+1):(\tilde{t}+k)}$
`exdqlmForecast` input	`start.t`	`k`	`fFF`	`fGG`

Table 4.3: Translation from mathematical symbols to additional parameters used in function `exdqlmForecast`.

in Section 4.4. For a full review of discount factors including discounting by component, see West et al. (1985).

4.2.4 k-step-ahead forecast

For an arbitrary time $\tilde{t}$, the k-step-ahead future marginal distribution of the quantile is

$$\mathbf{F}'_{\tilde{t}+k}\boldsymbol{\theta}_{\tilde{t}+k}|y_1,\ldots,y_{\tilde{t}} \sim \mathrm{N}(\mathbf{F}'_{\tilde{t}+k}a_{\tilde{t}}(k), \mathbf{F}'_{\tilde{t}+k}\mathbf{R}_{\tilde{t}}(k)\mathbf{F}_{\tilde{t}+k}) \tag{4.6}$$

where $\mathbf{a}_{\tilde{t}}(k) = \mathbf{G}_{\tilde{t}+k}\mathbf{a}_{\tilde{t}}(k-1)$, $\mathbf{R}_{\tilde{t}}(k) = \mathbf{G}_{\tilde{t}+k}\mathbf{R}_{\tilde{t}}(k-1)\mathbf{G}'_{\tilde{t}+k} + \mathbf{W}_{\tilde{t}+k}$, $\mathbf{a}_{\tilde{t}}(0) = \mathbf{m}_{\tilde{t}}$, and $\mathbf{R}_{\tilde{t}}(0) = \mathbf{C}_{\tilde{t}}$, with $\mathbf{m}_{\tilde{t}}$ and $\mathbf{C}_{\tilde{t}}$ denoting the filtered mean and covariance of $\boldsymbol{\theta}_{\tilde{t}}$, respectively. This distribution is implemented in the function `exdqlmForecast` which is compatible with the outputs from both the MCMC and ISVB algorithms. Table 4.3 shows the input parameters used in the function, as well as their mathematical symbols seen in this section. The output of `exdqlmForecast` includes the parameters of the forecasted distribution as well as $\mathbf{a}_{\tilde{t}}(k)$ and $\mathbf{R}_{\tilde{t}}(k)$ for the k steps.

4.2.5 Transfer function model

The transfer function extension of the exDQLM presented in Chapter 3 is also implemented within the **exdqlm** package, particularly for the ISVB algorithm. For readability, we briefly review the key details of this model here.

For $t = 1,\ldots,T$ and univariate input, X_t, a transfer function exDQLM with

exponential decay is defined by

$$y_t|\boldsymbol{\theta}_t, \gamma, \sigma \sim \text{exAL}_{p_0}(\mathbf{F}_t'\boldsymbol{\theta}_t + \zeta_t, \sigma, \gamma) \tag{4.7}$$

$$\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \mathbf{W}_t \sim \text{N}(\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t) \tag{4.8}$$

$$\zeta_t|\zeta_{t-1}, \psi_{t-1}, \omega_t \sim \text{N}(\lambda\zeta_{t-1} + X_t\psi_{t-1}, \omega_t) \tag{4.9}$$

$$\psi_t|\psi_{t-1}, \nu_t \sim \text{N}(\psi_{t-1}, \nu_t). \tag{4.10}$$

Here ζ_t captures the effect of the current and past regression effects on the quantile at time t. Alternatively, the parameter ψ_t determines the instantaneous effect X_t has on the quantile. The parameter $\lambda \in [0, 1]$ represents the memory of the regression effect up to time t. This effect decays at an exponential rate, reducing by a factor of λ at every time step. More explicitly, the effect of X_t on the quantile at time $t+k$ is $\lambda^k\psi_{t-1}X_t$. Thus, the series $k_t \geq \frac{\log(\epsilon)-\log(|\psi_{t-1}X_t|)}{\log(\lambda)}$ represents a lower bound for the number of time steps until the effect of X_t is less than or equal to a fixed ϵ. The median value of k_t corresponding to $\epsilon = 1\text{e}{-3}$ is included as an output of the transfer function routine.

Conditional on a fixed λ, say $\tilde{\lambda}$, we augment the ISVB algorithm to incorporate the transfer function structure by simply replacing $\mathbf{F}_t, \boldsymbol{\theta}_t, \mathbf{G}_t$, and $\mathbf{W}_t$ in Equations (4.1)-(4.2) with $\tilde{\mathbf{F}}_t' = (\mathbf{F}_t', 1, 0)$, $\tilde{\boldsymbol{\theta}}_t' = (\boldsymbol{\theta}_t', \zeta_t, \psi_t)$, $\tilde{\mathbf{G}}_t = \text{blockdiag}\left\{\mathbf{G}_t, \begin{pmatrix} \tilde{\lambda} & X_t \\ 0 & 1 \end{pmatrix}\right\}$, and $\tilde{\mathbf{W}}_t = \text{blockdiag}\left\{\mathbf{W}_t, \begin{pmatrix} \omega_t & 0 \\ 0 & \nu_t \end{pmatrix}\right\}$, respectively. This augmentation is implemented in the function `transfn_exdqlmISVB`, an extension of the routine found in `exdqlmISVB`.

Similar to the specification of $\mathbf{W_t}$ discussed in Section 4.2.3, discount factors $(\delta_\zeta, \delta_\psi)$ are utilized to specify $\begin{pmatrix} \omega_t & 0 \\ 0 & \nu_t \end{pmatrix}$. A bivariate normal conjugate prior on $(\zeta_0, \psi_0)' \sim N(\mathbf{m}_0^{tf}, \mathbf{C}_0^{tf})$ completes the transfer function model extension. Table 4.4 shows the additional parameters used in `transfn_exdqlmISVB`, as well as their

Symbol	X_t	$\tilde{\lambda}$	$(\delta_\zeta, \delta_\psi)$	$\mathbf{m}_0^{tf}$	$\mathbf{C}_0^{tf}$
`transfn_exdqlmISVB` input	`X`	`lam`	`tf.df`	`tf.m0`	`tf.C0`

Table 4.4: Translation from mathematical symbols to additional parameters used in function `transfn_exdqlmISVB`.

mathematical symbols discussed in this section.

4.2.6 Special cases

As mentioned previously, non-time-varying quantile regression is a special case of our methodology. More specifically, when discount factors are set to 1, the evolution covariance $\mathbf{W}_t$ reduces to $\mathbf{0}$ resulting in non-time-varying parameters (West et al., 1985). Thus, the methodology presented in Yan and Kottas (2017) which utilizes the exAL in a static model can also be implemented using the package with parameter `df = 1`.

The next special case of the exDQLM arises from the fact that the exAL is an extension of the AL (Yan and Kottas, 2017). When $\gamma = 0$, the observational error ϵ_t in Equation (4.1) reduces such that $\epsilon_t \sim \text{exAL}_{p_0}(\epsilon_t|0, \sigma, \gamma) = \text{AL}_{p_0}(\epsilon_t|0, \sigma)$. This special case with observational errors distributed independently from an AL is the dynamic quantile linear model (DQLM) presented in Gonçalves et al. (2020). The DQLM can be implemented with the package by fixing $\gamma = 0$ using the parameter `dqlm.ind = TRUE`.

The final special case is the intersection of the previous two. A static quantile regression model with observational errors distributed according to an AL is the well-known Bayesian parametric method for quantile regression first presented in Yu and Moyeed (2001). This methodology, utilized to create the Bayesian package for quantile regression **bayesQR**, can also be implemented with the package using parameters `df = 1` and `dqlm.ind = TRUE`.

4.3 Model Diagnostics

To evaluate the quantile inference and predictive performance of the exDQLM, the quantitative and visual model diagnostics presented in Section 2.2.3 are included in the package. We summarize them here.

The first diagnostic included in the package is one-step-ahead predictive distribution sequence. For the exDQLM, and conditional on the MAP estimates $\{\hat{v}_{1:T}, \hat{s}_{1:T}, \hat{\sigma}, \hat{\gamma}\}$, we estimate the sequence as

$$\hat{u}_t = \Phi(y_t | y_{1:t-1}, \hat{v}_{1:T}, \hat{s}_{1:T}, \hat{\sigma}, \hat{\gamma}) \tag{4.11}$$

where Φ denotes the normal CDF, $y_{1:t-1} = y_1, \ldots, y_{t-1}$, $v_{1:T} = (v_1, \ldots, v_T)$ and $s_{1:T} = (s_1, \ldots, s_T)$.

Using this estimated sequence $\{\hat{u}_t\}$, we can assess the model performance in several ways. First, we visually examine the correlation of the estimated sequence via an ACF plot. Second, by transforming the values with a standard normal inverse CDF, we compare their distributional shape to that of a standard normal with a normal QQ-plot. To quantify the divergence of this transformed sequence from normality we compute the KL divergence (Kullback and Leibler, 1951), $\text{KL}(h, \phi) = \int_{-\infty}^{\infty} h(x) \log \frac{h(x)}{\phi(x)} \text{d}x$, where h denotes the numerically approximated density of the transformed sample and ϕ is the standard normal density. A smaller KL divergence suggests a superior model fit. Lastly, the MAP standardized forecast errors,

$$\frac{y_t - \mathbb{E}[y_t | y_{1:t-1}, \hat{v}_{1:T}, \hat{s}_{1:T}, \hat{\sigma}, \hat{\gamma}]}{\sqrt{\text{Var}(y_t | y_{1:t-1}, \hat{v}_{1:T}, \hat{s}_{1:T}, \hat{\sigma}, \hat{\gamma})}} \tag{4.12}$$

can also be used as an additional graphical aid in examining the distribution at specific time points. These MAP standardized forecast errors are included in the

output of the `exdqlmMCMC` and `exdqlmISVB` functions.

For a final diagnostic, we include Gelfand and Ghosh (1998) posterior predictive loss criterion (pplc) with check loss function $\rho_{p_0}(x) = x[p_0 - I(x < 0)]$. That is,

$$\text{pplc} = \sum_t \mathbb{E}[\rho_{p_0}(y_t^{obs} - y_t^{rep})|y_{1:T}], \tag{4.13}$$

where the expectation is with respect to the posterior predictive distribution of y_t, $p(y_t^{rep}|y_{1:T})$. Again, a smaller pplc value suggest a superior model fit. Samples from the posterior replicate distributions are included in the output of the `exdqlmMCMC` and `exdqlmISVB` functions.

The function `exdqlmChecks` implements the model checking criteria discussed in this section. When an output from either `exdqlmMCMC` and `exdqlmISVB` is passed to `exdqlmChecks`, the function output includes: the estimated sequence $\{\hat{u}_t\}$, the KL divergence, pplc, the ordered pairs of the QQ-plot, and autocorrelations by lag. By default, the function also produces the QQ-plot, ACF plot and MAP standard forecast error plot. Examples of the utility of `exdqlmChecks` can be found in the following Section 4.4.

4.4 Examples

The following examples are implemented on a personal computer with a 2.5 GHz Intel Core i5 processor. We begin by loading the **exdqlm** package used for all examples.

```
R> library(exdqlm)
```

4.4.1 Lake Huron

In this example, we consider the Lake Huron time series from the package **data sets**. The data are annual measurements of the level (ft) of Lake Huron from 1875 to 1972, shown in Figure 4.2. With this example we show how to: build a basic state-space structure, specify a prior on γ, run the MCMC algorithm, plot the estimated quantiles and forecast distributions.

To estimate the dynamic distribution of the data, we consider three quantiles, $p_0 = 0.95, 0.50$, and 0.05. We begin by creating the state-space structure and prior used to estimate the quantiles (i.e. $\mathbf{F}_t, \mathbf{G}_t, \mathbf{m}_0$, and $\mathbf{C}_0$ discussed in Section 4.2). We choose to model the quantiles with a second order polynomial trend, which we construct with the `polytrendMod` function.

```
R> model = polytrendMod(order = 2, m0 = c(mean(LakeHuron),0),
+                       C0 = 10*diag(2))
R> model

$FF
[1] 1 0


$GG
     [,1] [,2]
[1,]    1    1
[2,]    0    1


$m0
[1] 579.0041   0.0000


$C0
```

```
     [,1] [,2]
[1,]   10    0
[2,]    0   10
```

To allow variability in the dynamic quantile, we select a single discount factor of 0.9 to define the evolution covariance matrix $\mathbf{W}_t$. Discount factors can also be optimized, an example of which we will show in Section 4.4.2. We use the parameter `fix.sigma = TRUE` to place a point-mass prior on σ to facilitate convergence, with point-mass at 0.07 for $p_0 = 0.95, 0.05$ and 0.4 for $p_0 = 0.5$. A discussion on how to choose the location of the point-masses can be found in Section 4.4.2. Further, we place truncated Student-t priors on the skewness parameters via `PriorGamma` as discussed in Section 4.2.1. Lastly, we limit the number of iterations with `n.burn = 700` and `n.mcmc = 300`. The R output is shown for only the last quantile using the input `verbose = TRUE`.

```
R> M95 = exdqlmMCMC(y = LakeHuron, p0 = 0.95, model = model,
+                   df = 0.9, dim.df = 2, fix.sigma = TRUE,
+                   sig.init = 0.07, PriorGamma =
+                        list(m_gam = -1, s_gam = 0.1, df_gam = 1),
+                   n.burn = 700, n.mcmc = 300, verbose = FALSE)
R> M50 = exdqlmMCMC(y = LakeHuron, p0 = 0.50, model = model,
+                    df = 0.9, dim.df = 2, fix.sigma = TRUE,
+                    sig.init = 0.4, PriorGamma =
+                        list(m_gam = 0, s_gam = 0.1, df_gam = 1),
+                    n.burn = 700, n.mcmc = 300, verbose = FALSE)
R> M5 = exdqlmMCMC(y = LakeHuron, p0 = 0.05, model = model,
+                   df = 0.9, dim.df = 2, fix.sigma = TRUE,
+                   sig.init = 0.07, PriorGamma =
```

```
+                         list(m_gam = 1, s_gam = 0.1, df_gam = 1),
+                    n.burn = 700, n.mcmc = 300, verbose = TRUE)

[1] "running isvb algorithm to initialize mcmc"
[1] "ISVB iteration 5: 2021-06-14 14:02:30"
[1] "ISVB iteration 10: 2021-06-14 14:02:31"
[1] "ISVB iteration 15: 2021-06-14 14:02:33"
[1] "ISVB converged: 15 iterations, 4.584 seconds"
[1] "burn-in iteration 500, acceptance rate 0.032:
    2021-06-14 14:03:11"
[1] "MCMC iteration 1000, acceptance rate 0.127: 2021-06-14 14:03:52"
[1] "MCMC complete: 1000 iterations, 78.617 seconds"
```

For this time series of length 98 it takes approximately 80 seconds for the MCMC algorithm sample 1000 iterations, as seen in the output above. By default `exdqlmMCMC` runs the ISVB algorithm to initialize the MCMC, which prints progress every 5 iterations. When the ISVB is used for initialization, the covariance matrix used to sample σ and γ in the MH step is automatically calculated from the ISVB samples of σ and γ. This covariance can be specified manually by using the setting `init.from.isvb = FALSE`, and providing the covariance via input `mh_Sig`. Further, using the setting `joint.sample = TRUE`, the covariance matrix can be recalculated from burned samples.

After the ISVB algorithm converges, the `exdqlmMCMC` prints progress after each 500 MCMC iterations, along with the current acceptance rate of the MH step and system time. It is not uncommon for low acceptance rates to facilitate convergence, particularly for extreme quantiles (i.e. 0.05), as is seen in this example. The data are relatively symmetric, therefore we do not expect the skewness parameter γ to be far from zero for $p_0 = 0.5$. We can examine this with the samples from the

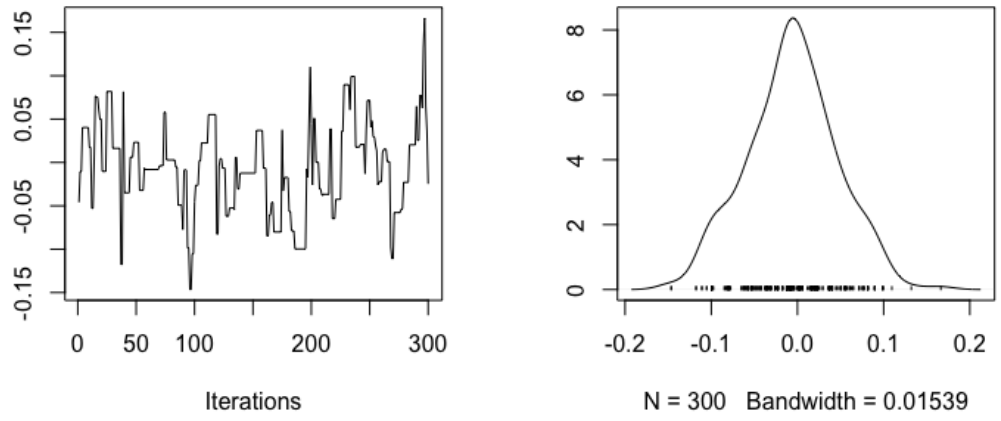

Figure 4.1: Example 1: Lake Huron. Trace plot (left) and density (right) of the MCMC samples of the skewness parameter γ estimated at the median.

posterior distribution of γ. All MCMC samples in the output of `exdqlmMCMC` are `mcmc` objects. Thus we can use the package **coda** to examine the trace plot and density of γ, seen in Figure 4.1.

```
R> library(coda)
R> traceplot(M50$samp.gamma, main = "")
R> densplot(M50$samp.gamma, main = "")
```

Clearly γ is indistinct from zero, thus the added flexibility of the skewness parameter is not necessary for the median. We re-run the model with a point-mass prior on γ at 0 using the settings `gam.init = 0` and `fix.gamma = 0`, equivalent to the setting `dqlm.ind = TRUE`.

```
R> M50 = exdqlmMCMC(y = LakeHuron, p0 = 0.50, model = model,
+                   df = 0.9, dim.df = 2,
+                   fix.sigma = TRUE, sig.init = 0.4,
```

```
+                   gam.init = 0, fix.gamma = TRUE,
+                   n.burn = 700, n.mcmc = 300)

[1] "running isvb algorithm to initialize mcmc"
[1] "ISVB iteration 5: 2021-07-18 12:42:47"
[1] "ISVB iteration 10: 2021-07-18 12:42:48"
[1] "ISVB iteration 15: 2021-07-18 12:42:50"
[1] "ISVB converged: 15 iterations, 4.305 seconds"
[1] "burn-in iteration 500, acceptance rate 0.212:
    2021-07-18 12:43:35"
[1] "MCMC iteration 1000, acceptance rate 0.265:
    2021-07-18 12:44:18"
[1] "MCMC complete: 1000 iterations, 87.92 seconds"
```

The results are `exdqlm` objects that can be used for analysis and forecasting. First we examine the estimated quantiles with the function `exdqlmPlot`, seen in Figure 4.2. This function generates a plot of the data with the MAP estimates and the 95% CrIs of the dynamic quantile. Subsequent quantiles (i.e. $p_0 = 0.50, 0.05$ in this example) can be added to the plot using the setting `add = TRUE`.

```
R> exdqlmPlot(y = LakeHuron, M95)
R> exdqlmPlot(y = LakeHuron, M50, add = TRUE, col = "blue")
R> exdqlmPlot(y = LakeHuron, M5, add = TRUE, col = "forest green")
R> legend("topright", lty = 1,
+           col = c("purple","blue","forest green"),
+           legend = c(expression('p'[0]*'=0.95'),
+                   expression('p'[0]*'=0.50'),
+                   expression('p'[0]*'=0.05')))
```

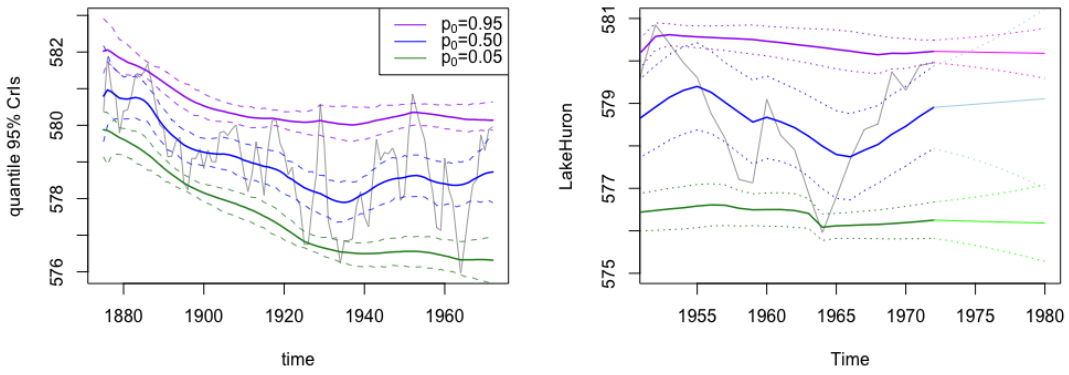

Figure 4.2: Example 1: Lake Huron. Left: MAP estimates and 95% CrIs of the estimated quantiles, plotted with the data (grey). Right: Forecasted quantile estimates and 95% credible intervals (seen after 1972), along with the filtered quantile estimates and 95% credible intervals (seen from 1952 to 1972).

Next, we forecast the $k = 8$ step-ahead distributions starting in 1972. To do this, we first create the observational vector (`fFF`) and evolution matrix (`fGG`) that will be used in the forecast updates. In this case both are non-time-varying and identical to those used to estimate the quantile.

```
R> fFF = model$FF
R> fGG =  model$GG
```

We plot the time series data in a narrower range for closer examination and add the forecast estimates using `exdqlmForecast` with setting `add = TRUE`. Notice we start the forecast at the last time index of the data, i.e. `start.t = length(LakeHuron)`.

```
R> plot.ts(LakeHuron,  xlim = c(1952,1980), ylim = c(575,581),
+        col = "dark grey")
R> exdqlmForecast(y = LakeHuron, start.t = length(LakeHuron),
+                k = 8, M95, fFF = fFF, fGG = fGG, plot = TRUE,
```

```
+                add = TRUE)
R> exdqlmForecast(y = LakeHuron, start.t = length(LakeHuron),
+                k = 8, M50, fFF = fFF, fGG = fGG, plot = TRUE,
+                add = TRUE, cols = c("blue","light blue"))
R> exdqlmForecast(y = LakeHuron, start.t = length(LakeHuron),
+                k = 8, M5, fFF = fFF, fGG = fGG, plot = TRUE,
+                add = TRUE, cols = c("forest green","green"))
```

This generates a plot of the data with the forecasted quantile estimates and 95% credible intervals (seen after 1972), along with the filtered quantile estimates and 95% credible intervals (before 1972), for reference. Results can be seen in Figure 4.2. The percentage in the CrIs can be modified with the parameter `cr.percent`.

4.4.2 Sunspots

For our next example, we will use the yearly Sunspot time series from the package **datasets**. The data are yearly counts for sunspots from 1700 to 1988, shown in Figure 4.3. With this example we will show how to: use the **dlm** package to create the state-space model, combine blocks of a state-space model, apply the ISVB algorithm, choose a prior for σ to facilitate convergence, perform visual model diagnostics to compare the exDQLM with the DQLM, and use those diagnostics for model selection.

The solar cycle impacts astronauts in space as well as lives and technology on earth. Of particular interest to scientists is the solar max, or period in which the sun is most active Fox (2012), therefore we will explore the 0.85 quantile in this example. Again, we begin by building the state-space structure used to estimate the quantile. Here we choose to model the quantile with a first order

polynomial trend component as well as a Fourier form seasonal component that incorporates a fundamental period every 11 observations (years), as well as some of its corresponding harmonics. This results in a dynamic model with a total of 9 state parameters, 1 for the first order polynomial component and 8 for the seasonal component. The **exdqlm** functions are compatible with time-invariant `dlm` objects from the **dlm** package. For example, we create the first order trend component with the function `dlmModPoly` from the **dlm** package.

```
R> library(dlm)
R> dlm.trend.comp = dlmModPoly(1, m0 = mean(sunspot.year), C0 = 10)
```

Next, we construct the seasonal component with the `seasMod` function. The duration of sunspot cycles typically follow a period of 11 years, of which we include the first 4 harmonics in the component.

```
R> seas.comp = seasMod(p = 11, h = 1:4, C0 = 10*diag(8))
```

To combine the two components into a single state-space structure, we use the function `combineMods`.

```
R> model = combineMods(dlm.trend.comp,seas.comp)
```

```
Warning message:
In combineMods(dlm.trend.comp, seas.comp) :
  m1 converted from a dlm object using 'dlmMod(m1)'
```

This produces a warning that the first input `dlm.trend.comp` was converted from a `dlm` object using the function `dlmMod`. This function can also be used directly on the `dlm` object to eliminate the warning. The resulting evolution matrix of the combined models is printed for illustration.

```
R> trend.comp = dlmMod(dlm.trend.comp)
R> model = combineMods(trend.comp,seas.comp)
R> model$GG
```

```
     [,1]    [,2]    [,3]    [,4]    [,5]    [,6]    [,7]    [,8]    [,9]
 [1,]    1  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
 [2,]    0  0.8413  0.5406  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
 [3,]    0 -0.5406  0.8413  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
 [4,]    0  0.0000  0.0000  0.4154  0.9096  0.0000  0.0000  0.0000  0.0000
 [5,]    0  0.0000  0.0000 -0.9096  0.4154  0.0000  0.0000  0.0000  0.0000
 [6,]    0  0.0000  0.0000  0.0000  0.0000 -0.1423  0.9898  0.0000  0.0000
 [7,]    0  0.0000  0.0000  0.0000  0.0000 -0.9898 -0.1423  0.0000  0.0000
 [8,]    0  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000 -0.6549  0.7557
 [9,]    0  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000 -0.7557 -0.6549
```

We will apply discount factors by component, i.e. a discount factor of 0.9 for the trend component of dimension 1 and a discount factor of 0.85 for the seasonal component of dimension 8. This discounting strategy can be applied with parameters `df = c(0.9,0.85)` and `dim.df = c(1,8)`.

Next we illustrate how to fix σ at a reasonable value. To do this, we investigate the scale of the data using the DQLM, a special case of the exDQLM described in Section 4.2.6. We call the function `exdqlmISVB` to apply the ISVB algorithm with `dqlm.ind = TRUE` and `fix.sigma = FALSE` to obtain estimates of the scale parameter σ. For this dataset of length 298 the ISVB algorithm for the DQLM takes 29 iterations and approximately 25 seconds to converge, seen in the output below. This output is suppressed for the remainder of this section using input `verbose = FALSE`.

```
R> M1 = exdqlmISVB(y = sunspot.year, p0 = 0.85, model = model,
```

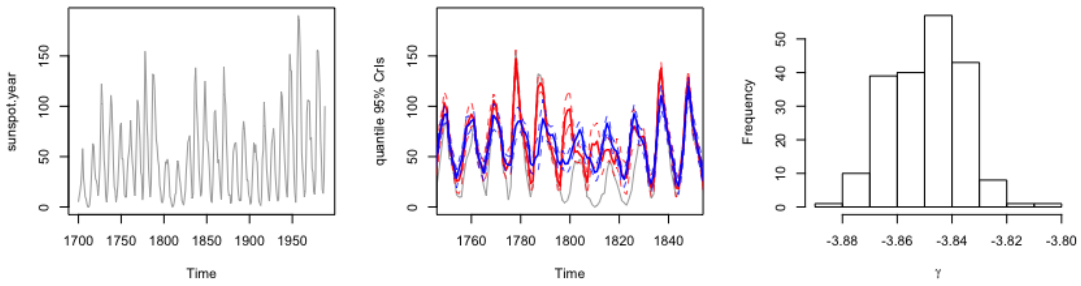

Figure 4.3: Example 2: Sunspots. Left: The sunspot time series from 1700 to 1988. Center: MAP estimates and 95% CrIs of the estimated quantiles from the DQLM (red) and exDQLM (blue), plotted with the data (grey) from 1750 to 1850. Right: Histogram of samples from the approximated posterior distribution of γ.

```
+                    df = c(0.9,0.85), dim.df = c(1,8),
+                    dqlm.ind = TRUE, fix.sigma = FALSE)

[1] "ISVB iteration 5: 2021-07-18 13:01:04"
[1] "ISVB iteration 10: 2021-07-18 13:01:08"
[1] "ISVB iteration 15: 2021-07-18 13:01:12"
[1] "ISVB iteration 20: 2021-07-18 13:01:16"
[1] "ISVB iteration 25: 2021-07-18 13:01:21"
[1] "ISVB converged: 29 iterations, 24.772 seconds"
```

The σ estimates from the DQLM give a general idea of the scale of the data and are a good starting point for fixing σ (or setting a strong prior) in the exDQLM. For example, the MAP estimate of σ from the DQLM is approximately 4.

```
R> summary(M1$samp.sigma)

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3.470   3.806   3.935   3.934   4.054   4.513
```

Reducing this value, we place a point-mass prior on σ at 2. By default, `exdqlmISVB` places a point-mass prior on σ at the value of `sig.init`. We re-run the algorithm

with `dqlm.ind = TRUE` to estimate the DQLM, as well as with the default settings to estimate the exDQLM.

```
R> M1 = exdqlmISVB(y = sunspot.year, p0 = 0.85, model = model,
+                  df = c(0.9,0.85), dim.df = c(1,8),
+                  dqlm.ind = TRUE, sig.init = 2,
+                  verbose = FALSE)
R> M2 = exdqlmISVB(y = sunspot.year, p0 = 0.85, model = model,
+                  df = c(0.9,0.85), dim.df = c(1,8),
+                  sig.init = 2, verbose = FALSE)
```

We can visualize the differences between the two resulting dynamic quantiles using the function `exdqlmPlot`, seen in Figure 4.3. For clarity we limit the figure to years 1750 to 1850.

```
R> plot(sunspot.year, xlim = c(1750,1850), col = "dark grey",
+       ylab = "quantile 95% CrIs")
R> exdqlmPlot(y = sunspot.year, M1, add = TRUE, col = "red")
R> exdqlmPlot(y = sunspot.year, M2, add = TRUE, col = "blue")
```

To examine whether the added flexibility of the exDQLM is significant, we can also visualize the approximate posterior distribution of the skewness parameter γ by plotting the samples from the corresponding variational distribution, also seen in Figure 4.3.

```
R> hist(M2$samp.gamma,xlab=expression(gamma),main="")
```

The approximate distribution is clearly distinct from zero, thus we expect a better model fit from the exDQLM.

To further examine the two models, we use the function `exdqlmChecks` to plot the model diagnostics discussed in Section 4.3. The results are seen in Figure

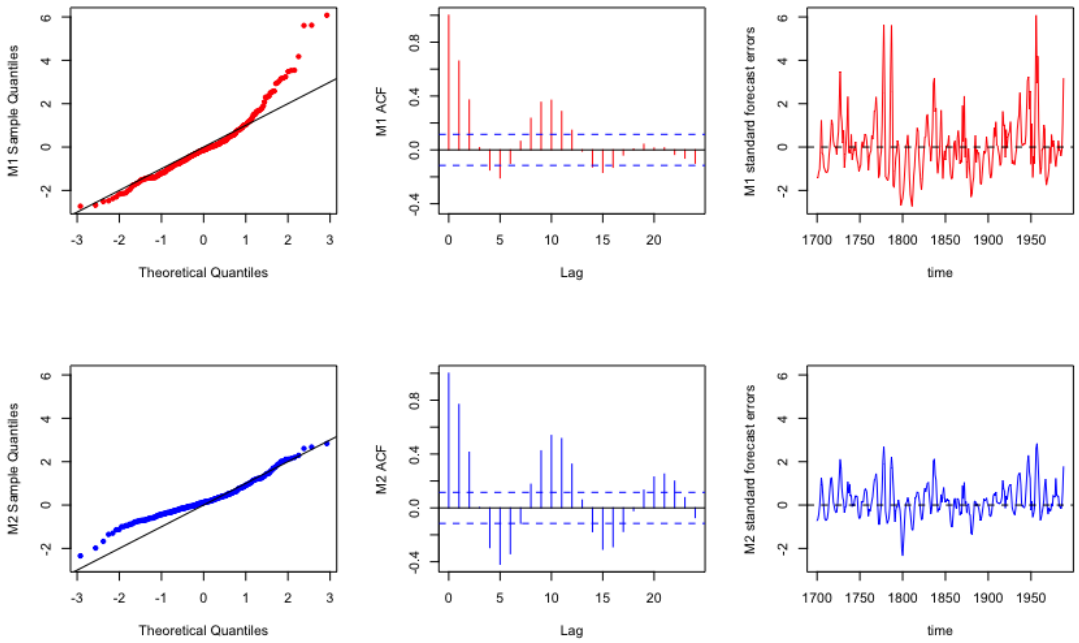

Figure 4.4: Example 2: Sunspots. The QQ-plots (left column), ACF plots (center column) and standard forecast errors (right column) from the DQLM (red) and exDQLM (blue).

4.4. It is clear the QQ-plot and the scale of the standard forecast errors favor the exDQLM.

```
R> exdqlmChecks(y = sunspot.year, M1, M2, cols = c("red","blue"))
```

Finally, the function `exdqlmChecks` can also be used for model selection. Say we want to check whether the discount factor we used for the seasonal component was optimal with respect to the KL divergence of the one-step-ahead predictive distributions. We can create a list of possible discount factors, quickly run the ISVB for each possible model, and select the set of discount factors which results in the lowest KL divergence. A simple example is seen below, which takes approximately 6 minutes to run.

```
R> possible.dfs = cbind(0.9,seq(0.85,1,0.05))
```

```
R> possible.dfs

     [,1] [,2]
[1,]  0.9 0.85
[2,]  0.9 0.90
[3,]  0.9 0.95
[4,]  0.9 1.00

R> KLs <- vector("numeric")
R> ref.samp = rnorm(length(sunspot.year))
R> for(i in 1:nrow(possible.dfs)){
+   temp.M2 = exdqlmISVB(y = sunspot.year, p0 = 0.85,
+               model = model, df = possible.dfs[i,],
+               dim.df = c(1,8), sig.init = 2, verbose = FALSE)
+   temp.check = exdqlmChecks(y = sunspot.year, temp.M2,
+                        plot = FALSE, ref = ref.samp)
+   KLs = c(KLs,temp.check$m1.KL)
+ }
R> # optimal dfs based off KL divergence
R> possible.dfs[which.min(KLs),]

[1] 0.90 0.85
```

The results suggest the discount factor we used for the seasonal component is optimal.

4.4.3 Big Tree water flow

For our final example, we consider the average monthly natural water flow (cubic feet per second) at the Big Tree gauge of the San Lorenzo river in Santa

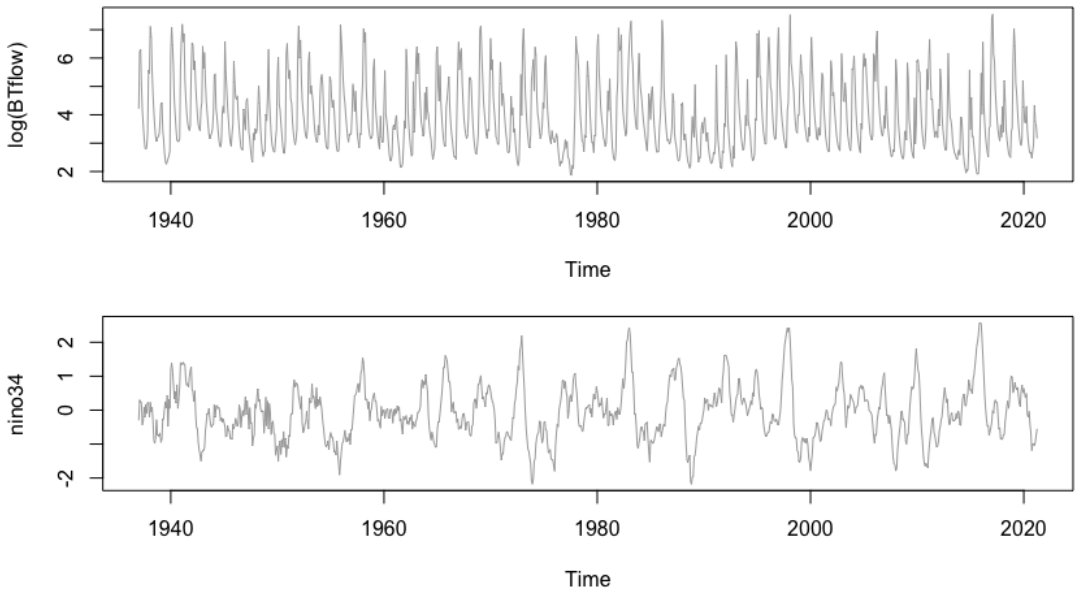

Figure 4.5: Example 3: Big Tree water flow. Top: Average monthly water flow at the Big Tree gauge of the San Lorenzo river in Santa Cruz County, CA, plotted on the log-scale. Bottom: Monthly values of Niño 3.4. Both time series span January 1937 through April 2021.

Cruz County, CA from January 1937 through April 2021 (U.S. Geological Survey, 2016), for a total of 1012 observations. The dataset is included in the package as `BTflow`. With this example we show how to build a state-space structure manually (particularly for a dynamic quantile regression), apply the transfer function exDQLM, examine individual components of the estimated quantiles, and perform quantitative model diagnostics to compare models.

For this particular dataset, the interest is in low quantiles as a possible indicator of drought, thus we will consider the $p_0 = 0.15$ quantile. To investigate any climatological influence on this low quantile, we include the Niño 3.4 index (Rayner et al., 2003) as in input of the exDQLM. The Niño 3.4 index is a measure of how anomalously warm or cool the average sea surface temperature is in the Niño 3.4 region (5°S-5°N, 170°W-120°W) of the Pacific Ocean. It is one of the most commonly used indices to define El Niño and La Niña events. The monthly

Niño 3.4 dataset used in the analysis is also included in the package as `nino34`. We model the relationship between Niño 3.4 and the quantile with two model structures: a dynamic quantile regression and a transfer function as seen in Section 4.2.5. To avoid modeling problems caused by the scale and extremes in the Big Tree dataset, we fit the following models to the time series on the log scale seen in Figure 4.5.

In addition to the components that capture the relationship with the Niño 3.4 index, we also include a first order polynomial trend component as well as a Fourier form seasonal component that includes a single periodic component with period 12. The Big Tree water flow time series follows a period of twelve months, of which we include only the first (annual) harmonic.

```
R> trend.comp = polytrendMod(1, m0 = 3, C0 = 0.1)
R> seas.comp = seasMod(p = 12, h = 1, C0 = diag(1,2))
R> model = combineMods(trend.comp, seas.comp)
```

Next, we create the dynamic regression component for the first model. A regression effect is easily written in a state-space format by setting the observational vector $F_t = X_t$ and $G_t = 1$, where X_t denotes the input at time t. Thus we create this regression component for the state-space model manually as follows.

```
R> reg.comp <- NULL
R> reg.comp$m0 = 0
R> reg.comp$C0 = 1
R> reg.comp$FF = matrix(nino34,nrow = 1)
R> reg.comp$GG = 1
```

Combining this with the trend and seasonal component we have the complete state-space structure for the dynamic quantile regression model.

```
R> model.w.reg = combineMods(model,reg.comp)
```

To allow the Niño 3.4 input components of both models to vary over time, we set the respective discount factors to be 0.95 (this is done via the parameter `df` in the `exdqlmISVB` function and the parameter `tf.df` in the `transfn_exdqlmISVB` function). Alternatively, we set the discount factors to 1 and 0.9 for the trend and seasonal components, respectively. This allows the seasonal amplitudes to evolve over time, while any further variation in the 0.15 quantile from year to year will solely be attributed to the effects of the Niño 3.4 input in the models.

The transfer function model is created within the function `transfn_exdqlmISVB` and follows the structure detailed in Section 4.2.5. Optimization using the KL divergence suggests the effects of the Niño 3.4 index decay at an optimal transfer function rate of $\lambda = 0.85$.

```
R> possible.lams = seq(0.1,0.9,0.05)
R> KLs <- vector("numeric")
R> ref.samp = rnorm(length(BTflow))
R> for(i in 1:length(possible.lams)){
R>   temp.M2 = transfn_exdqlmISVB(y = log(BTflow), p0 = 0.15,
+                         model = model, df = c(1,0.9),
+                         dim.df = c(1,2), X = nino34,
+                         tf.df = c(0.95), lam = possible.lams[i],
+                         tf.m0 = c(0,0), tf.C0 = diag(c(0.1,0.05)),
+                         sig.init = 0.1, gam.init = - 0.1,
+                         tol = 0.05, verbose = FALSE)
R>   temp.check = exdqlmChecks(y = log(BTflow), temp.M2,
+                      plot = FALSE, ref = ref.samp)
R>   KLs = c(KLs,temp.check$m1.KL)
```

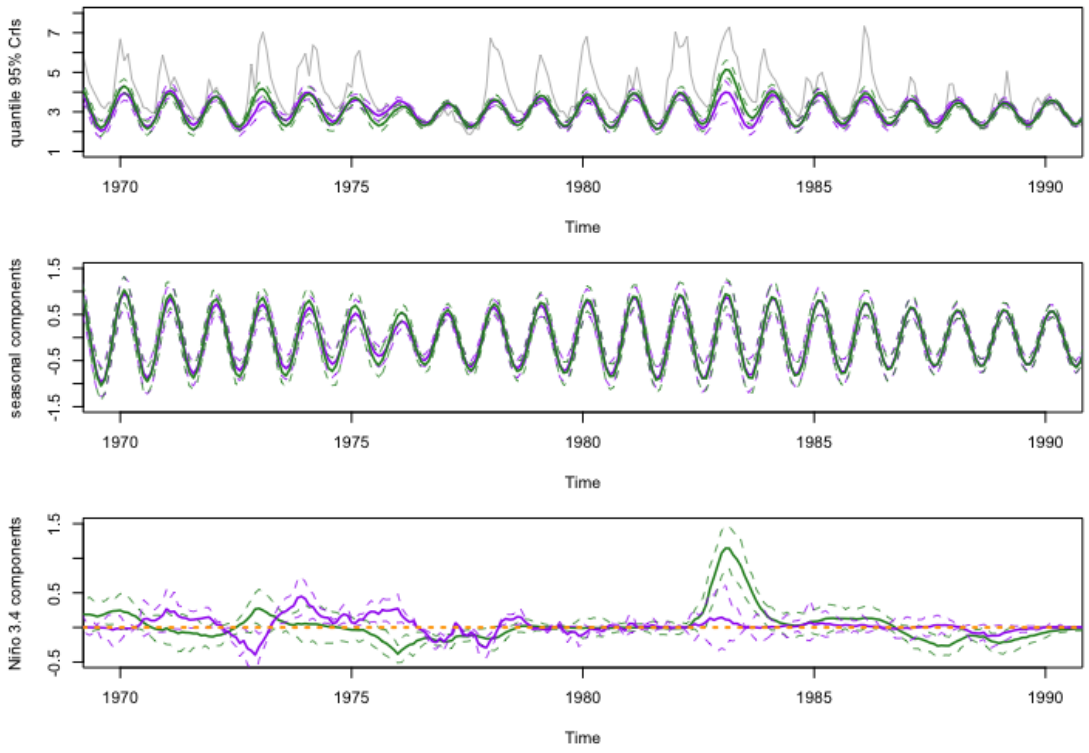

Figure 4.6: Example 3: Big Tree water flow. Top: MAP estimates and 95% CrIs of the estimated quantiles from the models with the regression component (purple) and transfer function component (green), plotted with the data (grey) from 1970 to 1990. Middle: MAP estimates and 95% CrIs of the annual seasonal components. Bottom: MAP estimates and 95% CrIs of the components which model the quantile association with Niño 3.4. The estimates of the dynamic regression component $X_t\boldsymbol{\theta}_{4,t}$ of M_1 where $\boldsymbol{\theta}_{4,t}$ denotes the fourth element of the state vector, are seen in purple. The estimates of the transfer function component ζ_t of M_2 are in green. The horizontal orange dotted line is at zero for reference.

```
R> }

R> # optimal lambda based off KL divergence

R> possible.lams[which.min(KLs)]

[1] 0.85
```

With these discount factors and optimal λ, we estimate the dynamic quantile regression model (`M1`) and transfer function model (`M2`).

```
R> M1 = exdqlmISVB(y = log(BTflow), p0 = 0.15, model = model.w.reg,
+                  df = c(1,0.9,0.95), dim.df = c(1,2,1),
+                  sig.init = 0.1, gam.init = - 0.1,
+                  tol = 0.05, verbose = FALSE)
R> M2 = transfn_exdqlmISVB(y = log(BTflow), p0 = 0.15,
+                     model = model, df = c(1,0.9), dim.df = c(1,2),
+                     X = nino34, tf.df = c(0.95), lam = 0.85,
+                     tf.m0 = c(0,0), tf.C0 = diag(c(0.1,0.005),2),
+                     sig.init = 0.1, gam.init = - 0.1,
+                     tol = 0.05, verbose = FALSE)
```

To visualize the differences between the two resulting dynamic quantiles we use the function `exdqlmPlot`, results of which are found in Figure 4.6. For clarity we limit the figure to view 1970 to 1990.

```
R> plot(log(BTflow), col = "grey", ylim = c(1,8),
+       xlim = c(1970,1990), ylab = "quantile 95% CrIs")
R> exdqlmPlot(y = BTflow, M1, add = TRUE)
R> exdqlmPlot(y = BTflow, M2, add = TRUE, col = "forest green")
```

Clearly the resulting quantiles have significant differences. To identify from where these differences arise, we can utilize the function `compPlot` to examine the components of the models more closely. We examine the annual component of the parameter vector first. The annual component is comprised of the second and third elements of the parameter vector $\boldsymbol{\theta}_t$, thus we set the parameter `index = c(2,3)` to visualize this component. The result is seen in Figure 4.6, again limited to the years 1970 to 1990.

```
R> plot(NA, ylim = c(-1.5,1.5), xlim = c(1970,1990),
+         ylab = "seasonal components")
```

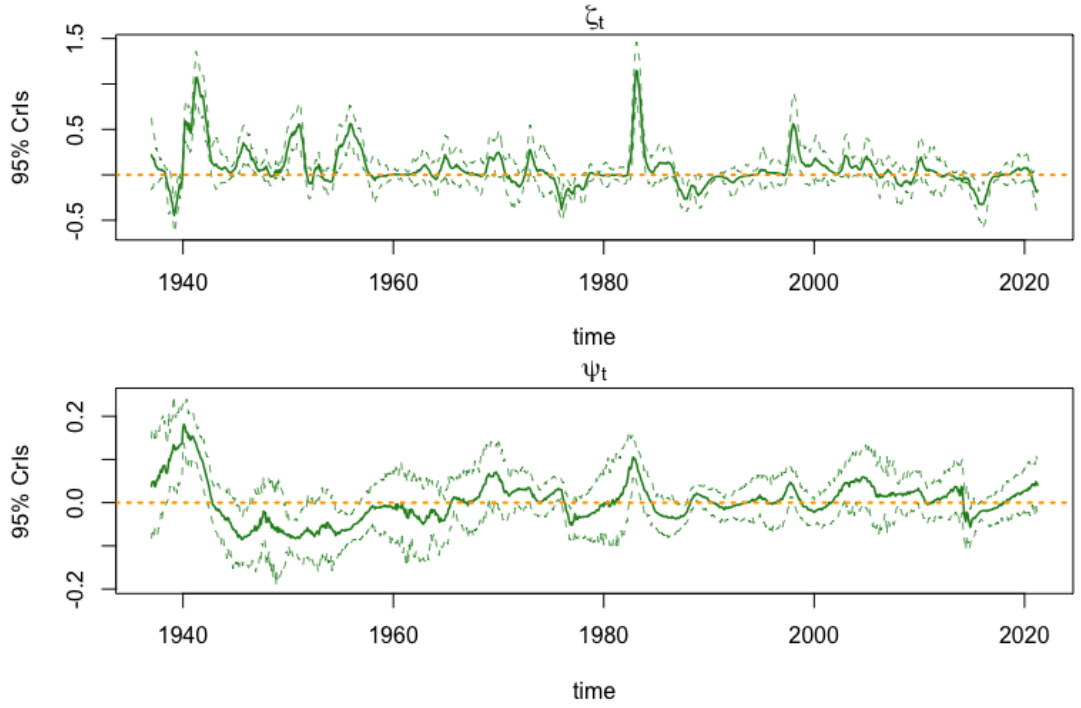

Figure 4.7: Example 3: Big Tree water flow. MAP estimates and 95% CrIs of parameters ζ_t and ψ_t from the transfer function component. Horizontal dotted orange lines are at zero for reference.

```
R> compPlot(y = log(BTflow), M1, index = c(2,3), add = TRUE)
R> compPlot(y = log(BTflow), M2, index = c(2,3), add = TRUE,
+            col = "forest green")
```

Although the seasonal components of the two models exhibit differences in amplitude, we expect more significant differences to appear in the Niño 3.4 component of each model. Again, we use the function `compPlot` to visualize these components seen in Figure 4.6 from 1970 to 1990.

```
R> plot(NA, ylim = c(-0.5,1.5), xlim = c(1970,1990),
+          ylab = "Niño 3.4 components")
R> compPlot(y = log(BTflow), M1, index = c(4), add = TRUE)
R> compPlot(y = log(BTflow), M2, index = c(4,5), add = TRUE,
```

```
+             col = "forest green")
R> abline(h = 0, col = "orange", lty = 3, lwd = 2)
```

For the transfer function model, `M2`, it is also of interest to look at the estimates of parameters ζ_t and ψ_t directly. Because these parameters augmented the state-space model as discussed in Section 4.2.5, the function `compPlot` can also be used to look at these estimates directly using the parameter `just.theta = TRUE`. This setting returns the dynamic posterior estimates of a single index of the state vector $\boldsymbol{\theta}_t$ (or $\tilde{\boldsymbol{\theta}}_t$ in the case of the transfer function model). Parameter `index` of `4` and `5` correspond to ζ_t and ψ_t, respectively. Results can be found in Figure 4.7.

```
R> compPlot(y = log(BTflow), M2, index = 4, col= "forest green",
+             add = FALSE, just.theta = TRUE)
R> abline(h = 0, col = "orange", lty = 3, lwd = 2)
R> title(expression(eta[t]))
R> compPlot(y = log(BTflow), M2, index = 5, col = "forest green",
+             add = FALSE, just.theta = TRUE)
R> abline(h = 0, col = "orange", lty = 3, lwd = 2)
R> title(expression(psi[t]))
```

To further understand these transfer function parameters, we examine the median number of time steps until the effect of Niño 3.4 on the 0.15 quantile is less than or equal to 1e−3, as discussed in Section 4.2.5. A median value 20.76 suggests the effect of the Niño 3.4 index value at time t last for around 21 months.

```
R> M2$median.kt

[1] 20.76305
```

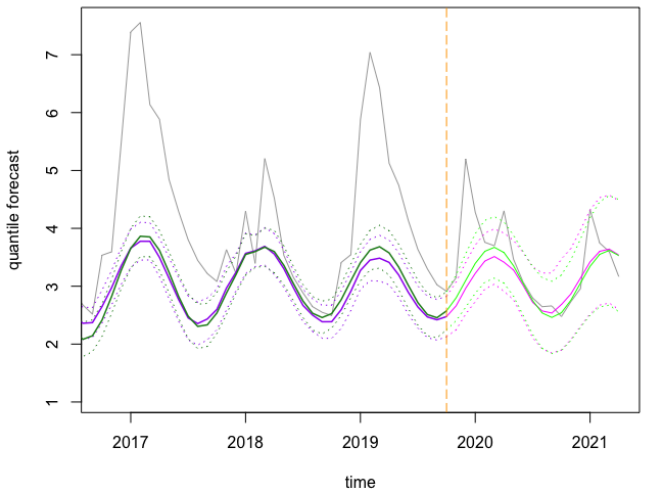

Figure 4.8: Example 3: Big Tree water flow. 18-step-ahead quantile forecast beginning October 2019 through April 2021 overlaid on the log Big Tree water flow time series. The vertical dot-dashed line is at the beginning of the forecast period, October 2019, for reference. Solid lines indicate mean estimates and dotted lines 95% CrI, with the filtered estimates seen leading up to October 2019 and the forecast estimates beyond October 2019. Estimates from M_1, the quantile regression model, are seen in purple/pink. Estimates from M_2, the transfer function model, are seen in green.

This suggests a long-term forecast could be of interest. Thus we forecast the $k = 18$ step ahead distributions for both models starting in October 2019, results of which are seen in Figure 4.8.

```
R> exdqlmForecast(y = log(BTflow), start.t = length(BTflow)-18,
+                   k = 18, M1)
R> exdqlmForecast(y = log(BTflow), start.t = length(BTflow)-18,
+                   k = 18, M2, add = TRUE,
+                   cols = c("forest green","green"))
```

Finally, we use the function `exdqlmChecks` to compute the KL divergence and pplc discussed in Section 4.3, results of which are seen in Table 4.5.

```
R> checks.out = exdqlmChecks(y = log(BTflow), M1, M2,
```

	M1		M2	
	output	value	output	value
pplc	`m1.pplc`	957.72	`m2.pplc`	897.22
KL	`m1.kl`	1.135	`m2.kl`	0.880

Table 4.5: Example 3: Big Tree water flow. Model diagnostic output from `exdqlmChecks` and resulting values.

```
+                          col = c("purple","orange"), plot = FALSE)
```

The KL divergence and pplc both clearly favor the nonlinear relationship captured by the transfer function model.

4.5 Conclusion

The exDQLM framework presented in the previous chapters provides a robust and versatile tool for dynamic quantile modeling, implemented in the presented **exdqlm** package. Relevant features of the data can easily be incorporated in the evolution of the quantile. The MCMC algorithm provides a method for exact posterior inference. Alternatively, the ISVB algorithm provides a method for fast approximate inference enabling estimation for large data sets, quick examination relevant model features and model selection. Further, the transfer function model provides a straight-forward method for quantifying a nonlinear relationship between a quantile and response. Finally, the forecast and diagnostic routines facilitate further model exploration. These capabilities have been discussed in detail and demonstrated with a set of examples. The **exdqlm** package makes these methods, and more generally, dynamic quantile modeling, accessible to all levels of users.

Chapter 5

Joint Estimation of Time-Varying Quantiles via Flexible Multivariate Dynamic Models

In this chapter, we develop a multivariate extended dynamic quantile linear model to consider multiple time series simultaneously and jointly estimate a specified quantile for each series. To do this, we first develop a multivariate exAL distribution. We then present the details of multivariate MCMC and ISVB algorithms for exact and approximate posterior inference, respectively. The utility of the multivariate model is illustrated via application to two real datasets.

5.1 Background

We begin by presenting a new class of multivariate distributions, the multivariate exAL (MVexAL). This family of distributions is an extension of the exAL and the multivariate AL (MVAL) developed in Kotz et al. (2003) which we review here.

Let $\mathbf{y} = (y_1, \ldots, y_n)'$ denote a vector of n observations distributed according to a MVAL, i.e. $\mathbf{y} \sim \text{MVAL}_{\mathbf{p}}(\boldsymbol{\mu}, \boldsymbol{\sigma})$. Kotz et al. (2003) establish that the MVAL is equal in distribution to the mixture representation,

$$\mathbf{y} = \boldsymbol{\mu} + A(\mathbf{p}) \odot \mathbf{v} + \sqrt{B(\mathbf{p}) \odot \mathbf{v} \odot \boldsymbol{\sigma}} \odot \mathbf{z} \tag{5.1}$$

with $\odot$ denoting element-wise multiplication. Here $\boldsymbol{\sigma} = (\sigma_1, \ldots \sigma_n)'$ is a set of scale parameters, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$ is a set of location parameters, and $\mathbf{p} = (p_1, \ldots, p_n)'$ are skewness parameters which will correspond to the quantiles of interest for the n observations. $A(\mathbf{p})$ and $B(\mathbf{p})$ are element-wise scalar functions analogous to those seen in Chapter 2 2.1, i.e. $A(\mathbf{p}) = \frac{1-2\mathbf{p}}{\mathbf{p}(1-\mathbf{p})}$ and $B(\mathbf{p}) = \frac{2}{\mathbf{p}(1-\mathbf{p})}$. Further, $\mathbf{v} = (v_1, \ldots v_n)'$ is such that each element v_j is marginally distributed according to an exponential distribution with scale σ_j for $j = 1, \ldots, n$. Finally, $\mathbf{z} = (z_1, \ldots z_n)' \sim \text{N}_n(\mathbf{0}, \mathbf{R})$, independent of $\mathbf{v}$, with $\mathbf{R}$ denoting a latent correlation matrix of dimension n.

Kotz et al. (2003) showed that each component y_j of a MVAL random vector $\mathbf{y}$ admits the representation

$$y_j = \mu_j + A(p_j)v_j + \sqrt{B(p_j)v_j\sigma_j}z_j \tag{5.2}$$

where z_j is a standard normal variable. This is the mixture representation of the univariate AL presented in Section 2.1, thus each element of $\mathbf{y}$ marginally follows an $\text{AL}_{p_j}(\mu_j, \sigma_j)$. For further details on the development and properties of the MVAL, see Kotz et al. (2003).

5.1.1 The multivariate exAL

We expand the advantages of the exAL to the multivariate setting by extending the multivariate AL presented in Kotz et al. (2003). Let $\mathbf{p}_0 = (p_{01}, \ldots, p_{0n})'$ such that each element p_{0j} denotes the quantile of interest for the j^{th} observation of $\mathbf{y}$, for $j = 1, \ldots, n$. Utilizing the mixture representation of the MVAL discussed above, we define a multivariate exAL (MVexAL) for inference of the $\mathbf{p}_0$-quantiles as follows:

$$\mathbf{y} = \boldsymbol{\mu} + C(\mathbf{p}, \boldsymbol{\gamma}) \odot \boldsymbol{\sigma} \odot |\boldsymbol{\gamma}| \odot \mathbf{s} + A(\mathbf{p}) \odot \mathbf{v} + \sqrt{B(\mathbf{p}) \odot \mathbf{v} \odot \boldsymbol{\sigma}} \odot \mathbf{z}. \tag{5.3}$$

Analogous to the univariate exAL, $\boldsymbol{\gamma} = (\gamma_1, \ldots \gamma_n)'$ is a set of skewness parameters and $\mathbf{s} = (s_1, \ldots s_n)'$ is such that each element of $\mathbf{s}$ is marginally distributed according to a standard half-normal distribution, independent of $\mathbf{v}$ and $\mathbf{z}$. The parameter $\mathbf{p}$ is now a scalar function of $\mathbf{p}_0$ and $\boldsymbol{\gamma}$; $\mathbf{p} = p(\mathbf{p}_0, \boldsymbol{\gamma}) = I(\boldsymbol{\gamma} < 0) + \{[\mathbf{p}_0 - I(\boldsymbol{\gamma} < 0)]/g(\boldsymbol{\gamma})\}$ where $g(\boldsymbol{\gamma}) = 2\Phi(-|\boldsymbol{\gamma}|)\exp(\boldsymbol{\gamma}^2/2)$ and $\Phi(\cdot)$ denotes the standard normal CDF also applied element-wise. Each element γ_j of the parameter vector $\boldsymbol{\gamma}$ has bounded support over the interval (L_j, U_j) where L_j is the negative root of $g(\gamma_j) = 1 - p_{0j}$ and U_j is the positive root of $g(\gamma_j) = p_{0j}$. Further, $A(\mathbf{p})$, $B(\mathbf{p})$, $C(\mathbf{p}, \boldsymbol{\gamma})$ are scalar functions of $\mathbf{p}$ with $A(\mathbf{p})$, $B(\mathbf{p})$ as defined in Equation (5.1) and $C(\mathbf{p}, \boldsymbol{\gamma}) = [I(\boldsymbol{\gamma} > 0) - \mathbf{p}]^{-1}$. Kindred to the univariate exAL, when $\boldsymbol{\gamma} = \mathbf{0}$ the MVexAL reduces to the MVAL.

Each component y_j of the MVexAL random vector $\mathbf{y}$ admits the representation

$$y_j = \mu_j + C(p_j, \gamma_j)\sigma_j|\gamma_j|s_j + A(p_j)v_j + \sqrt{B(p_j)v_j\sigma_j}z_j \tag{5.4}$$

where again z_j is a standard normal variable. Clearly this is the mixture representation of the univariate exAL presented in Section 2.1, thus each element of $\mathbf{y}$

marginally follows an $\text{exAL}_{p_j}(\mu_j, \sigma_j, \gamma_j)$. Thus, by construction μ_j corresponds to the fixed quantile p_{0j} such that

$$\int_{-\infty}^{\mu_j} \text{exAL}(y_j|\mu_j, \sigma_j, \gamma_j)\text{d}y_j = p_{0j}. \tag{5.5}$$

As in the univariate setting, the capacity to fix the quantiles of interest is retained. Further, the mixture representation preserves many of the computational advantages discussed in Chapter 2, making the MVexAL a robust error distribution for our multivariate dynamic quantile model.

5.1.2 Specifying s and v

common v & common s	$\frac{C_iC_j\lvert\gamma_i\rvert\lvert\gamma_j\rvert[1-4\phi(0)^2]+A_iA_j+\sqrt{B_iB_j}\pi R_{ij}}{4\sqrt{(C_i^2\gamma_i^2[1-4\phi(0)^2]+A_i^2+B_i)(C_j^2\gamma_j^2[1-4\phi(0)^2]+A_j^2+B_j)}}$
common v & iid s_j	$\frac{A_iA_j+\sqrt{B_iB_j}\pi R_{ij}}{4\sqrt{(C_i^2\gamma_i^2[1-4\phi(0)^2]+A_i^2+B_i)(C_j^2\gamma_j^2[1-4\phi(0)^2]+A_j^2+B_j)}}$
iid v_j & common s	$\frac{C_iC_j\lvert\gamma_i\rvert\lvert\gamma_j\rvert[1-4\phi(0)^2]+\sqrt{B_iB_j}\pi R_{ij}}{4\sqrt{(C_i^2\gamma_i^2[1-4\phi(0)^2]+A_i^2+B_i)(C_j^2\gamma_j^2[1-4\phi(0)^2]+A_j^2+B_j)}}$
iid v_j & iid s_j	$\frac{\sqrt{B_iB_j}\pi R_{ij}}{4\sqrt{(C_i^2\gamma_i^2[1-4\phi(0)^2]+A_i^2+B_i)(C_j^2\gamma_j^2[1-4\phi(0)^2]+A_j^2+B_j)}}$

Table 5.1: $\text{Corr}(y_i, y_j)$ for the four possible modeling options of latent parameters v_j and s_j.

As mentioned previously, for $j = 1, \ldots, n$ the latent parameters v_j and s_j are marginally distributed according to an exponential with mean σ_j and a standard half normal, respectively. Here we discuss four modeling options for the elements of $\mathbf{v}$ and $\mathbf{s}$: (a) common v and common s for all $j = 1, \ldots, n$, (b) common v and iid s_j, (c) iid v_j and common v, and (d) iid v_j and iid s_j.

To examine the possible behaviors of the correlation of the multivariate model at various quantiles, we derive $\text{Corr}(y_i, y_j)$ for the four possible cases, seen in Table 5.1. We use the short-hand $A_j = A(p_j)$, $B_j = B(p_j)$, and $C_j = C(p_j, \gamma_j)$. Further,

R_{ij} denotes the ij-th element of the latent correlation matrix $\mathbf{R}$. Note, in cases (a) and (b), common v also implies common σ for $j = 1, \ldots, n$. This implication is too restrictive of an assumption for our applications, thus we focus on the other two case (c) and (d) going forward.

The correlations in Table 5.1 depend on the parameters $\boldsymbol{\gamma}$ and fixed quantiles $\mathbf{p}_0$ via the functions A_j, B_j, and C_j. To investigate further the behavior of the correlation for the remaining cases, we plot the correlations for three quantiles of varying extremes and two differing values of the correlation R_{ij}. It can be seen in Figure 5.1 that for iid v_j and common s, when the skewness parameters γ_i and γ_j are distinct from zero, the correlations tend to 1 for the extreme quantiles regardless of the values of R_{ij}. In contrast, for the case of iid v_j and iid s_j seen in Figure 5.2, the correlations weaken at the extreme quantiles, particularly for skewness parameters distinct from zero. Although this feature limits the amount of correlation that can be captured by the model, the flexibility for $\boldsymbol{\gamma}$ distinct from zero is desirable. Thus, due to the increasing correlations at the extreme quantiles for case (c) with common s, we utilize case (d) with both iid s_j and v_j for $j = 1, \ldots n$ going forward.

5.2 The MVexDQLM

Consider a set of n dimensional time-evolving responses, $\mathbf{y}_t = (y_{t,1}, \ldots, y_{t,n})'$, for times $t = 1, \ldots, T$. For each t, a general multivariate dynamic model can be defined by

$$\text{Observation equation:} \qquad \mathbf{y}_t = \mathbf{F}_t'\boldsymbol{\theta}_t + \boldsymbol{\epsilon}_t \tag{5.6}$$

$$\text{System Equation:} \qquad \boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t. \tag{5.7}$$

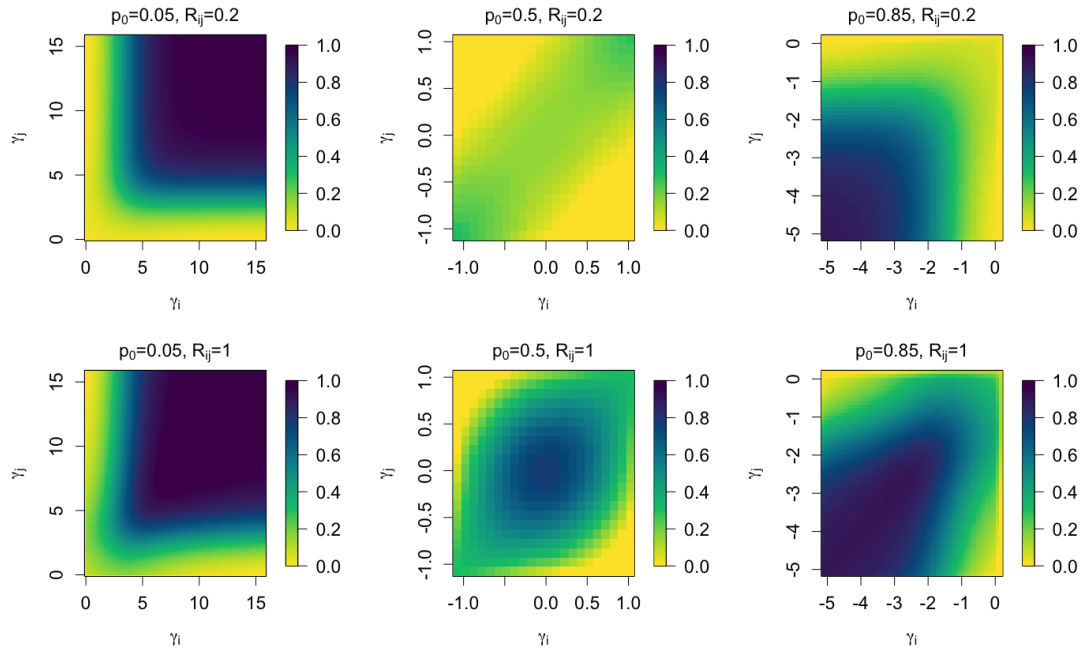

Figure 5.1: Corr(y_i, y_j) for case (c); common s and iid v_j, for varying values of γ_i (x-axes), γ_j (y-axes), p_0 (columns) and R_{ij} (rows).

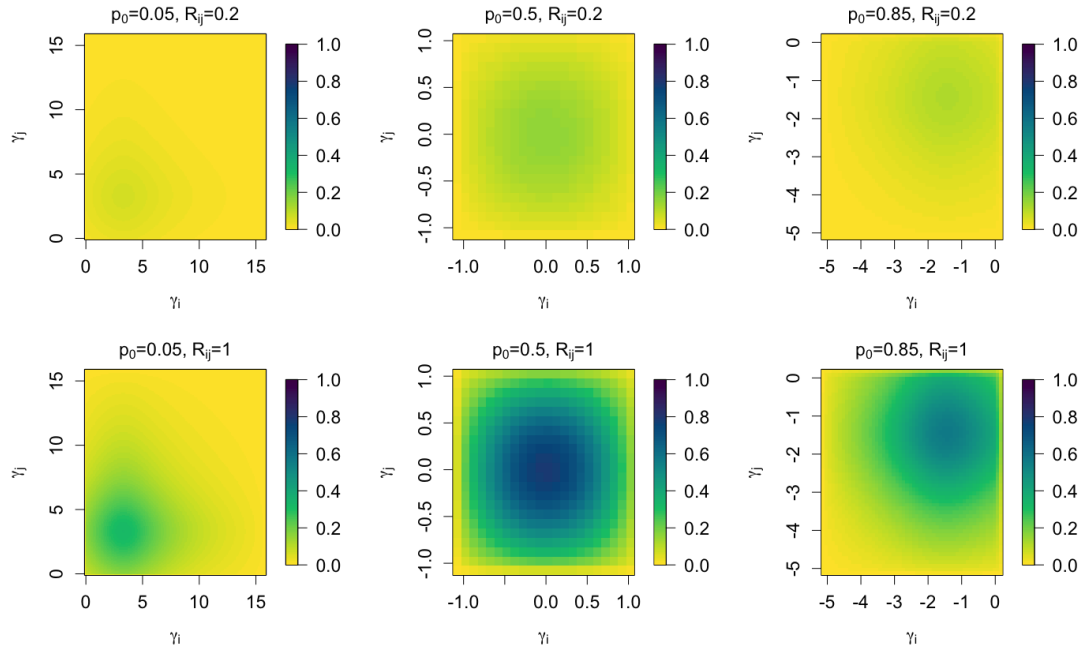

Figure 5.2: Corr(y_i, y_j) for case (d); iid s_j and iid v_j, for varying values of γ_i (x-axes), γ_j (y-axes), p_0 (columns) and R_{ij} (rows).

Here $\mathbf{F}_t$ is the $q \times n$ regression matrix of the covariates corresponding to the $q \times 1$ regression parameter vector $\boldsymbol{\theta}_t$ at time t, and $\mathbf{G}_t$ is the q-dimensional evolution matrix defining the structure of the parameter vector evolution in time. We propose a multivariate extended dynamic quantile linear model (MVexDQLM) for inference on the $\mathbf{p}_0 = (p_{01}, \ldots, p_{0n})'$ quantiles of the elements of $\mathbf{y}_t$ by specifying the observational errors of a multivariate dynamic model to be distributed independently from the MVexAL, which we denote $\text{MVexAL}_{\mathbf{p}_0}$. Utilizing a Gaussian time-evolving structure on the system error vector, $\boldsymbol{\omega}_t \sim \text{N}(\mathbf{0}, \mathbf{W}_t)$ where $\mathbf{W}_t$ is the evolution variance matrix, our MVexDQLM can be written

$$\mathbf{y}_t | \boldsymbol{\theta}_t, \boldsymbol{\gamma}, \boldsymbol{\sigma} \sim \text{MVexAL}_{\mathbf{p}_0}(\mathbf{F}_t'\boldsymbol{\theta}_t, \boldsymbol{\sigma}, \boldsymbol{\gamma}) \tag{5.8}$$

$$\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}, \mathbf{W}_t \sim \text{N}(\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t) \tag{5.9}$$

where the normal distribution according to which $\boldsymbol{\theta}_t$ evolves is q-variate. The specific structure of $\{\mathbf{F}_t, \boldsymbol{\theta}_t, \mathbf{G}_t, \mathbf{W}_t\}$ will depend heavily on the application, thus this is discussed in more detail within Sections 5.4 and 5.5. Similar to the exDQLM, the mixture representation of the MVexAL in Equation (5.3) can be exploited to rewrite the MVexDQLM as the following hierarchical model for $t = 1, \ldots, T$:

$$\mathbf{y}_t | \boldsymbol{\theta}_t, \boldsymbol{\sigma}, \boldsymbol{\gamma}, \mathbf{v}_t, \mathbf{s}_t \sim \tag{5.10}$$

$$\text{N}(\mathbf{F}_t'\boldsymbol{\theta}_t + \mathbf{C} \odot \boldsymbol{\sigma} \odot |\boldsymbol{\gamma}| \odot \mathbf{s}_t + \mathbf{A} \odot \mathbf{v}_t, \sqrt{\mathbf{B} \odot \mathbf{v}_t \odot \boldsymbol{\sigma}}' \mathbf{R} \sqrt{\mathbf{B} \odot \mathbf{v}_t \odot \boldsymbol{\sigma}})$$

$$v_{t,j}, s_{t,j} | \sigma_j \sim \text{Exp}(v_{t,j} | \sigma_j) \text{N}^+(s_{t,j} | 0, 1), \quad \text{for } j = 1, \ldots, n \tag{5.11}$$

$$\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}, \mathbf{W}_t \sim \text{N}(\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t). \tag{5.12}$$

where the normal distribution according to which $\mathbf{y}_t$ evolves is n-variate. Here, $\mathbf{A} = A(\mathbf{p})$, $\mathbf{B} = B(\mathbf{p})$, $\mathbf{C} = C(\mathbf{p}, \boldsymbol{\gamma})$ are the functions of $\mathbf{p}$ and $\boldsymbol{\gamma}$ defined in Equa-

tion (5.3). Under the Bayesian framework, we place priors on the parameters as follows: a q-variate prior $\boldsymbol{\theta}_0 \sim \text{N}(\mathbf{m}_0, \mathbf{C}_0)$, $\sigma_j \sim \text{IG}(a_\sigma, b_\sigma)$ and $\gamma_j \sim t_{(L_j,U_j)}(0,1)$ for $j = 1, \ldots, n$. As in Chapter 2, we utilize discount factors to specify the covariance matrix $\mathbf{W}_t$. Note that if a point mass prior at zero is used for skewness parameters $\boldsymbol{\gamma}$, the observational errors $\boldsymbol{\epsilon}_t$ in Equation (5.6) are distributed independently from the MVAL, a special case of the MVexAL defined in Equation (5.1). This simplified model for fixed quantiles $\mathbf{p} = \mathbf{p}_0$, we call the MVDQLM. However, since we already explored in detail the added advantages of the exAL in comparison to the AL in Chapters 2 and 3, we will not be exploring the MVDQLM further in this Chapter.

5.2.1 MCMC algorithm

The mixture representation of the MVexAL facilitates Bayesian posterior simulation using a MCMC algorithm. Closed form conditional distributions are available for the latent variables $\mathbf{s}_{1:T} = \{\mathbf{s}_1, \ldots, \mathbf{s}_T\}$ which are sampled with Gibbs steps. In the multivariate setting we loose conjugacy for the latent variables $\mathbf{v}_{1:T} = \{\mathbf{v}_1, \ldots, \mathbf{v}_T\}$ and scale parameters $\boldsymbol{\sigma}$. Thus, $\mathbf{v}_{1:T}$, $\boldsymbol{\sigma}$, skewness parameters $\boldsymbol{\gamma}$ and correlation matrix $\mathbf{R}$ are updated with MH steps. Conditional on $\mathbf{s}_{1:T}$, $\mathbf{v}_{1:T}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\gamma}$, the dynamic regression coefficients can be sampled using a forward filtering backwards sampling (FFBS), full details of which can be found in Appendix A. MCMC posterior simulation is summarized in Algorithm 3.

5.2.2 ISVB algorithm

The high computation and memory costs seen with the exDQLM are compounded in the multivariate setting, thus an efficient algorithm for accurate estimation is even more necessary. We extend our ISVB algorithm to the multivariate

Algorithm 3: MVexDQLM MCMC

Initialize $\boldsymbol{\sigma}^{(0)}, \boldsymbol{\gamma}^{(0)}, \mathbf{v}_{1:T}^{(0)}, \mathbf{s}_{1:T}^{(0)}, \boldsymbol{\theta}_{1:T}^{(0)}$;

for *i = 0, . . . , I-1* **do**

1. Sample $(\boldsymbol{\gamma}^{(i+1)}, \boldsymbol{\sigma}^{(i+1)})|\boldsymbol{\theta}_{1:T}^{(i)}, \mathbf{v}_{1:T}^{(i)}, \mathbf{s}_{1:T}^{(i)}, \mathbf{R}^{(i)}$ using a Metropolis-Hastings step with a Gaussian random walk with joint proposals on the logit and log scales.

2. Sample $\mathbf{R}^{(i+1)}|\boldsymbol{\theta}_t^{(i)}, \mathbf{v}_{1:T}^{(i)}, \mathbf{s}_{1:T}^{(i)}, \boldsymbol{\sigma}^{(i)}, \boldsymbol{\gamma}^{(i)}$ using a Metropolis-Hastings step with a Rescaled-Wishart proposal distribution centered at $\mathbf{R}^{(i)}$.

3. **for** *t=1,. . . ,T* **do**

Sample $\mathbf{v}_t^{(i+1)}|\boldsymbol{\theta}_t^{(i)}, \mathbf{s}_t^{(i)}, \boldsymbol{\sigma}^{(i)}, \boldsymbol{\gamma}^{(i)}, \mathbf{R}^{(i)}$ using a Metropolis-Hastings step with a Gaussian random walk proposal on the log scale.

end

4. **for** *t=1,. . . ,T* **do**

Sample $\mathbf{s}_t^{(i)}|\boldsymbol{\theta}_t^{(i)}, \mathbf{v}_t^{(i)}, \boldsymbol{\sigma}^{(i)}, \boldsymbol{\gamma}^{(i)}, \mathbf{R}^{(i)} \sim \text{MVN}^{+}(\boldsymbol{\mu}_{\mathbf{s}_t}, \boldsymbol{\sigma}^2_{\mathbf{s}_t})$, where MVN^{+} denotes a multivariate normal truncated to the positive reals with parameters

$$\Sigma_{\mathbf{s}_t} = \left[\text{diag}\left\{ \frac{\mathbf{C}^{(i)} \odot \boldsymbol{\gamma}^{(i)} \odot \sqrt{\boldsymbol{\sigma}^{(i)}}}{\sqrt{\mathbf{B}^{(i)} \odot \mathbf{v}_t^{(i)}}} \right\} \mathbf{R}^{-1} \text{diag}\left\{ \frac{\mathbf{C}^{(i)} \odot \boldsymbol{\gamma}^{(i)} \odot \sqrt{\boldsymbol{\sigma}^{(i)}}}{\sqrt{\mathbf{B}^{(i)} \odot \mathbf{v}_t^{(i)}}} \right\} + \mathbf{I}_n \right]^{-1},$$

$$\boldsymbol{\mu}_{\mathbf{s}_t} = \Sigma_{\mathbf{s}_t} \text{diag}\left\{ \frac{\mathbf{C}^{(i)} \odot \boldsymbol{\gamma}^{(i)} \odot \sqrt{\boldsymbol{\sigma}^{(i)}}}{\sqrt{\mathbf{B}^{(i)} \odot \mathbf{v}_t^{(i)}}} \right\} \mathbf{R}^{-1} \text{diag}\left\{ \frac{1}{\sqrt{\mathbf{B}^{(i)} \odot \mathbf{v}_t^{(i)} \odot \boldsymbol{\sigma}^{(i)}}} \right\} (\mathbf{y}_t - \mathbf{F}_t' \boldsymbol{\theta}_t^{(i)} - \mathbf{A}^{(i)} \odot \mathbf{v}_t^{(i)}).$$

end

5. **for** *t=1,. . . ,T* **do**

Sample $\boldsymbol{\theta}_t|\mathbf{v}_{1:T}^{(i)}, \mathbf{s}_{1:T}^{(i)}, \boldsymbol{\gamma}^{(i)}, \boldsymbol{\sigma}^{(i)}, \mathbf{R}^{(i)}$ via FFBS. The forward part of the FFBS algorithm uses the forecast distribution $p(\mathbf{y}_t|\mathbf{D}_{t-1}) = \text{N}(\mathbf{f}_t, \mathbf{Q}_t)$ where $\mathbf{D}_{t-1} = \{\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}\}$,

$$\mathbf{f}_t = \mathbf{F}_t' \mathbf{a}_t + \mathbf{C}^{(i)} \odot \boldsymbol{\sigma}^{(i)} \odot |\boldsymbol{\gamma}^{(i)}| \odot \mathbf{s}_t^{(i)} + \mathbf{A}^{(i)} \odot \mathbf{v}_t^{(i)}, \quad \mathbf{Q}_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + \boldsymbol{\sigma}^{(i)} \odot \mathbf{B}^{(i)} \odot \mathbf{v}_t^{(i)}.$$

end

end

setting as follows.

Let $\boldsymbol{\xi} = \{\boldsymbol{\theta}_{1:T}, \boldsymbol{\sigma}, \boldsymbol{\gamma}, \mathbf{v}_{1:T}, \mathbf{s}_{1:T}, \mathbf{R}\}$ denote the set of all parameters in the MVexDQLM. We approximate the posterior distribution $f(\boldsymbol{\xi}|\mathbf{y}_{1:T})$ with an arbitrary variational distribution $r(\boldsymbol{\xi})$ which minimizes the KL divergence and equivalently the ELBO, as detailed in Chapter 2 Section 2.2.2. We utilize a mean-field

approximation with the following factorization,

$$r(\boldsymbol{\xi}) = r(\boldsymbol{\theta}_{1:T})r(\mathbf{v}_{1:T})r(\mathbf{s}_{1:T})r(\boldsymbol{\sigma}, \boldsymbol{\gamma})r(\mathbf{R}). \tag{5.13}$$

The ELBO is maximized by the following

$$r(\boldsymbol{\xi}_c) \propto \exp\left\{\int \log f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{-c})r(\boldsymbol{\xi}_{-c})\mathrm{d}\boldsymbol{\xi}_{-c}\right\} \tag{5.14}$$

where $\boldsymbol{\xi}_c$ denotes the set of variables in the component being maximized and $\boldsymbol{\xi}_{-c}$ denotes the variables not in that component of the partition. We initialize the partitioned variational distributions seen in Equation (5.13) and iteratively update the distributions using Equation (5.14) until convergence. For the MVexDQLM, the variational distributions $r(\mathbf{s}_{1:T})$ and $r(\boldsymbol{\theta}_{1:T})$ are recognizable, closed-form distributions which can be updated directly at each iteration. In particular, $r(\boldsymbol{\theta}_{1:T})$ is updated using a FFBSm algorithm. Alternatively, variation distributions $r(\mathbf{v}_{1:T})$, $r(\boldsymbol{\sigma}, \boldsymbol{\gamma})$, and $r(\mathbf{R})$ are not recognizable and therefore updated at each iteration using adaptive IS (AIS). ISVB posterior inference for the MVexDQLM is summarized in Algorithm 4 with details of the FFBSm and IS algorithms found in Appendix A.

5.3 Comparison criteria

To evaluate the predictive ability of the MVexDQLM, we again consider the Gelfand and Ghosh (1998) pplc with check loss function ρ_{p_0}. Given the posterior replicate distribution of $\mathbf{y}_t$, $p(\mathbf{y}_t^{rep}|\mathbf{D}_T)$, the pplc for the MVexDQLM can be

Algorithm 4: MVexDQLM ISVB part 1

Set $k = 0$ and initialize $r^0(\mathbf{s}_{1:T})$, $r^0(\mathbf{v}_{1:T})$, $r^0(\boldsymbol{\theta}_0)$, $r^0(\boldsymbol{\sigma}, \boldsymbol{\gamma})$ and $r^0(\mathbf{R})$;
while convergence has not been achieved

1. **for** t=1,...,T **do**

Update $r^{(k+1)}(\mathbf{s}_t) = \text{MVN}^+({\boldsymbol{\Sigma}_{s_t}}^{(k+1)}\boldsymbol{\mu}_{s_t}^{(k+1)}, {\boldsymbol{\Sigma}_{s_t}}^{(k+1)})$, where the ij-th element of $\boldsymbol{\Sigma}_{s_t}^{-1(k+1)}$ for $i \neq j$ is

$$\left(\boldsymbol{\Sigma}_{s_t}^{-1(k+1)}\right)_{ij} = \left\langle \frac{C_i\sqrt{\sigma_i}\gamma_i C_j\sqrt{\sigma_j}\gamma_j}{\sqrt{B_i B_j}} \right\rangle^{(k)} \left\langle \frac{1}{\sqrt{v_{ti}v_{tj}}} \right\rangle^{(k)} \langle R^{-1}\rangle_{ij}^{(k)},$$

for $i, j = 1, \ldots, n$, where γ_i, σ_i, B_i and C_i denote the i-th elements of $\boldsymbol{\gamma}$, $\boldsymbol{\sigma}$, $\mathbf{B}$ and $\mathbf{C}$, respectively, and $\langle R^{-1}\rangle_{ij}$ denotes the ij-th element of $\langle R^{-1}\rangle$. For $i = j$,

$$\left(\boldsymbol{\Sigma}_{s_t}^{-1(k+1)}\right)_{ii} = \left\langle \frac{C_i^2\sigma_i\gamma_i^2}{B_j} \right\rangle^{(k)} \left\langle \frac{1}{v_{ti}} \right\rangle^{(k)} + 1,$$

for $i = 1, \ldots, n$. Further, the i-th element of $\boldsymbol{\mu}_{s_t}^{(k+1)}$ is

$$\begin{aligned}\left(\boldsymbol{\mu}_{s_t}^{(k+1)}\right)_i = &\left\langle \frac{C_i|\gamma_i|}{B_i} \right\rangle^{(k)} \left\langle \frac{1}{v_{ti}} \right\rangle^{(k)} (y_{ti} - \langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_i^{(k)})\langle R^{-1}\rangle_{ii}^{(k)} \\ &+ \sum_{j=1, j\neq i}^{n} \left\langle \frac{C_j|\gamma_j|\sigma_j}{\sqrt{B_j B_i \sigma_i}} \right\rangle^{(k)} \left\langle \frac{1}{\sqrt{v_{tj}v_{ti}}} \right\rangle^{(k)} (y_{tj} - \langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_j^{(k)})\langle R^{-1}\rangle_{ij}^{(k)} \\ &+ \left\langle \frac{C_i|\gamma_i|A_i}{B_i} \right\rangle^{(k)} \langle R^{-1}\rangle_{ii}^{(k)} \\ &+ \sum_{j=1, j\neq i}^{n} \left\langle \frac{C_j|\gamma_j|\sigma_j A_j}{\sqrt{B_j B_i \sigma_i}} \right\rangle^{(k)} \left\langle \sqrt{\frac{v_{ti}}{v_{tj}}} \right\rangle^{(k)} \langle R^{-1}\rangle_{ij}^{(k)}\end{aligned}$$

where $\langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_i$ and A_i denote the i-th elements of $\langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle$ and $\mathbf{A}$, respectively, for $i = 1, \ldots, n$.

end

2. **for** t=1,...,T **do**

Update the smoothed distribution $r^{(k+1)}(\boldsymbol{\theta}_t) = \text{N}(\mathbf{m}_t^s, \mathbf{C}_t^s)$ using a FFBSm algorithm with forecast distribution $r^{(k+1)}(\mathbf{y}_t|\mathbf{D}_{t-1}) = \text{N}(\mathbf{f}_t, \mathbf{Q}_t)$ where $\mathbf{f}_t = \mathbf{F}_t'\mathbf{a}_t + \tilde{\mathbf{Q}}_t\tilde{\mathbf{f}}_t$ and $\mathbf{Q}_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + \tilde{\mathbf{Q}}_t$. Here, the i-th element of $\tilde{\mathbf{f}}$ is

$$\begin{aligned}\tilde{\mathbf{f}}_i = &\left\langle \frac{C_i|\gamma_i|}{B_i} \right\rangle^{(k)} \left\langle \frac{1}{v_{ti}} \right\rangle^{(k)} \langle s_{ti}\rangle^{(k)}\langle R^{-1}\rangle_{ii}^{(k)} \\ &+ \sum_{j=1, j\neq i}^{n} \left\langle \frac{C_j|\gamma_j|\sqrt{\sigma_j}}{\sqrt{B_j B_i \sigma_i}} \right\rangle^{(k)} \left\langle \frac{1}{\sqrt{v_{tj}v_{ti}}} \right\rangle^{(k)} \langle s_{tj}\rangle^{(k)}\langle R^{-1}\rangle_{ij}^{(k)}\end{aligned}$$

and the ij-th element of $\tilde{\mathbf{Q}}_t^{-1}$ for $i \neq j$ is

$$(\tilde{\mathbf{Q}}_t^{-1})_{ij} = \left\langle \frac{1}{\sqrt{\sigma_i\sigma_j B_i B_j}} \right\rangle^{(k)} \left\langle \frac{1}{\sqrt{v_{ti}v_{tj}}} \right\rangle^{(k)} \langle R^{-1}\rangle_{ij}^{(k)}$$

and for $i = j$

$$(\tilde{\mathbf{Q}}_t^{-1})_{ii} = \left\langle \frac{1}{\sigma_i B_i} \right\rangle^{(k)} \left\langle \frac{1}{v_{ti}} \right\rangle^{(k)} \langle R^{-1}\rangle_{ii}^{(k)}.$$

end

Algorithm 5: MVexDQLM ISVB part 2

3. **for** $t=1,\ldots,T$ **do**
Update $r^{(k+1)}(\mathbf{v}_t)$ using AIS with proposal distribution

$$\text{MVT}^{+}(\langle \mathbf{v}_t \rangle^{(k)}, \text{Cov}(\mathbf{v}_t)^{(k)})$$

where MVT^{+} denotes a multivariate Student-t distributions truncated to the positive reals with 1 degree of freedom and $\text{Cov}(\mathbf{v}_t)^{(k)} = \langle (\mathbf{v}_t - \langle \mathbf{v}_t \rangle^{(k)})(\mathbf{v}_t - \langle \mathbf{v}_t \rangle^{(k)})' \rangle^{(k)}$.

end

4. Update $r^{(k+1)}(\boldsymbol{\sigma}, \boldsymbol{\gamma})$ using AIS with proposal distributions $\text{MVT}_{(\mathbf{L},\mathbf{U})}(\langle \boldsymbol{\gamma} \rangle^{(k)}, \text{Cov}(\boldsymbol{\gamma})^{(k)})$ and $\text{MVT}^{+}(\langle \boldsymbol{\sigma} \rangle^{(k)}, \text{Cov}(\boldsymbol{\sigma})^{(k)}))$ for $\boldsymbol{\gamma}$ and $\boldsymbol{\sigma}$, respectively, where $\text{MVT}_{(\mathbf{L},\mathbf{U})}$ denotes a multivariate Student-t distribution truncated to the n-dimensional support $(\mathbf{L}, \mathbf{U})$.

5. Update $r^{(k+1)}(\mathbf{R})$ using AIS with proposal distribution Rescaled-Wishart$(\nu, \langle R \rangle^{(k)})$. Further details of this AIS and the other AIS steps can be found in Appendix A

6. Set $k = k + 1$.

end

computed as

$$\text{pplc} = \sum_j \sum_t \mathbb{E}[\rho_{p_0}(\mathbf{y}_{t,j}^{obs} - \mathbf{y}_{t,j}^{rep}) | \mathbf{D}_T]. \tag{5.15}$$

Further, we consider the Rosenblatt (1952) one-step-ahead predictive distribution function for each component of the time-series, as suggested in Prado et al. (2006). That is, for $j = 1, \ldots, n$, $\{u_{t,j}\}$ will be a sequence of independent random variables uniformly distributed on $(0, 1)$. For the MVexDQLM,

$$u_{t,j} = \Pr(Y_{tj} \leq y_{tj} | \mathbf{D}_{t-1}, \boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}) = \Phi(y_{tj} | \mathbf{D}_{t-1}, \boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}). \tag{5.16}$$

Conditional on $\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}$, the predictive distribution of $\mathbf{y}_t$ is normally distributed with mean $\mathbf{f}_t$ and variance $\mathbf{Q}_t$ seen in Algorithms 3 and 4, thus $u_{tj} = \Phi(y_{tj} | f_{tj}, Q_{tj})$ where f_{tj} denotes the j-th element of $\mathbf{f}_t$ and Q_{tj} denotes the j-th diagonal element of $\mathbf{Q}_t$. We can obtain estimates for u_{tj} conditionally on MAP estimates of $\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}$ from the MCMC or ISVB posterior samples. For each component of the time-

series, we can diagnose the model performance using the distributional shape of the estimated sequence $\{\hat{u}_{t,j}\}$ through transformation with a standard normal inverse CDF, as done in Chapter 2 Section 2.2.3. This will allow visual diagnosis with a normal QQ-plot, as well as quantitative diagnosis using the KL divergence.

5.4 Big Tree water flow analysis

Water flow levels act as useful indicators for potential floods as well as droughts. In particular, percentiles of historical stream flow are used locally as as drought severity classifications (Lombard, 2016; NOAA National Centers for Environmental Information, 2021). In this example, we examine the monthly water flow (cubic feet per second; cfs) at the Big Tree (BT) water gauge on the San Lorenzo river in Santa Cruz county, CA. In particular, we estimate a low quantile as a possible indicator of below-normal water levels in Santa Cruz county.

Many other climatological and local weather variables are also known to affect water supply. Two of these such variables are precipitation and temperature, which we also include in this analysis. In particular, when rainfall is significantly low for an extended period, water levels decrease (Johannis et al., 2016). Further, rising temperatures can cause enhanced evaporation and early snow melts resulting in reduced surface water (Ostroff et al., 2017). To incorporate these factors into our analysis, we estimate a low quantile of the monthly precipitation (inches; in) as well as a high quantile of temperature (Fahrenheit; F), both recorded in Santa Cruz county.

Not only are water flow, precipitation and temperature known indicators of extreme water conditions, they are also known to be practically associated with one another. Thus, this example is an ideal application for the MVexDQLM which provides a framework for not only estimating the quantiles, but also incorporating

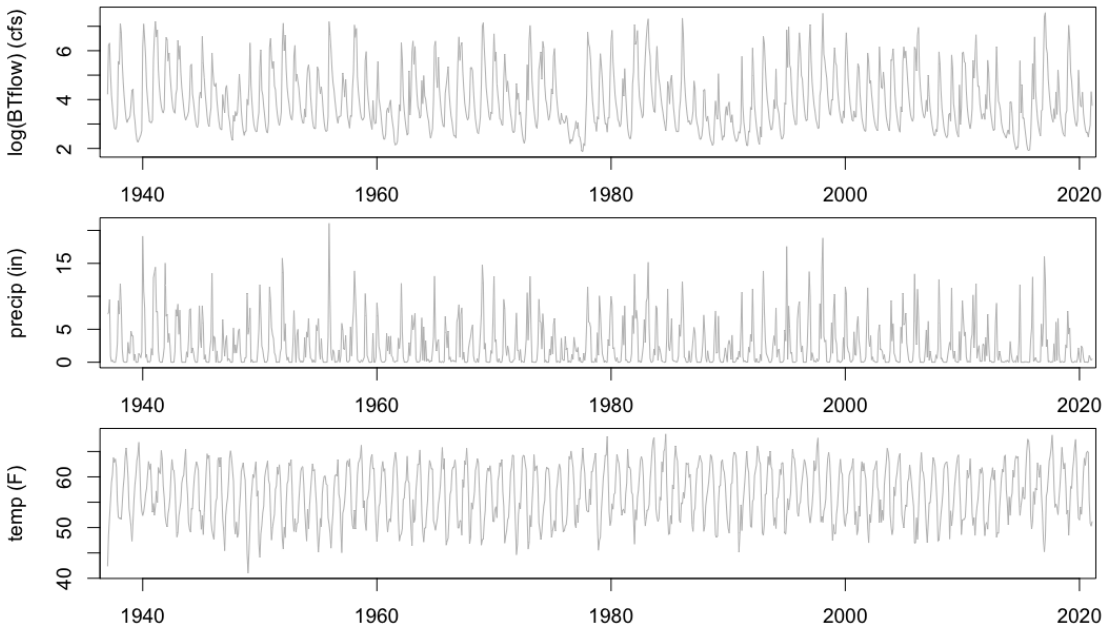

Figure 5.3: Top panel: Monthly water flow on the log-scale (cfs) at the Big Tree water gauge on the San Lorenzo river in Santa Cruz county, CA. Middle panel: Monthly precipitation (in) recorded in Santa Cruz county. Bottom panel: Monthly temperature (F) also recorded in Santa Cruz county. All data span from January 1937 to February 2021.

any correlation structure between them. The time-series used in the analysis span from January 1937 to February 2021, and can be seen in Figure 5.3. To avoid modeling problems caused by the scale and extremes in the BT water flow dataset, we fit the following models to the BT time-series on the log scale.

5.4.1 Modeling framework

As discussed above, in this example the interest is in quantiles as possible indicators of drought. For both water flow and precipitation, we examine the 0.15-quantile as abnormally low values suggest dry conditions for these time-series. Alternatively, drought conditions are correlated with high temperatures, therefore we consider the 0.85-quantile for the temperature time-series. Thus, we set $\mathbf{p}_0 = (0.15, 0.15, 0.85)'$ for this analysis.

To assess the gain of including a multivariate structure between the time-series, we fit two separate models to estimate the $\mathbf{p}_0$-quantiles. The first, M_0, is the univariate version of the second, M_1. M_0 is equivalent to the exDQLM presented in Chapter 2, i.e. the special case of the MVexDQLM with latent correlation $\mathbf{R} = \mathbf{I}_n$ where $\mathbf{I}_n$ denotes an identity matrix of dimension n. Alternatively, M_1 estimates the multivariate structure between the time-series via the latent correlation matrix $\mathbf{R}$.

We begin by developing the evolution structure of the quantiles. For each time-series, we select significant components as well as optimize discount factors in the univariate setting. In particular, for the Big Tree water flow time-series, we model the 0.15-quantile with a first order baseline component, η_t^{BT}, as well as a Fourier form seasonal component that includes a single periodic component with period of twelve months, $\alpha_{1,t}^{\text{BT}(1)}$, the first (annual) harmonic. Following the notation defined in Chapter 2 Section 3.3.1, these components for the Big Tree water flow quantile can be represented as

$$\boldsymbol{\theta}_t^{\text{BT}} = (\eta_t^{\text{BT}}, \alpha_{1,t}^{\text{BT}(1)}, \alpha_{2,t}^{\text{BT}(1)})' \tag{5.17}$$

$$\mathbf{F}_t^{\text{BT}} = (1, 1, 0)' \tag{5.18}$$

$$\mathbf{G}_t^{\text{BT}} = \text{blockdiag}(1, \mathbf{G}_{\alpha,1}^{\text{BT}}). \tag{5.19}$$

where $\mathbf{G}_{\alpha,l}$ denotes the lth seasonal evolution matrix, i.e. $\mathbf{G}_{\alpha,l} = \begin{pmatrix} \cos(\frac{2\pi}{12}l) & \sin(\frac{2\pi}{12}l) \\ -\sin(\frac{2\pi}{12}l) & \cos(\frac{2\pi}{12}l) \end{pmatrix}$. Further, we specify the evolution covariance, $\mathbf{W}_t^{\text{BT}}$, using the two discount factors $\boldsymbol{\delta}^{\text{BT}} = (0.86, 0.82)$ for the baseline and seasonal components, respectively. Finally, we fix $\sigma_{\text{BT}} = 0.1$ to facilitate convergence.

Next, for the precipitation time-series, we model the 0.15-quantile with a first order baseline component, η_t^{prec}, as well as a Fourier form seasonal component that includes the first two harmonics, $\alpha_{1,t}^{\text{prec}(l)}$ for $l = 1, 2$, the annual and semi-annual

harmonics, respectively. These components for the precipitation quantile can be represented as

$$\boldsymbol{\theta}_t^{\text{prec}} = (\eta_t^{\text{prec}}, \alpha_{1,t}^{\text{prec}(1)}, \alpha_{2,t}^{\text{prec}(1)}, \alpha_{1,t}^{\text{prec}(2)}, \alpha_{2,t}^{\text{prec}(2)})' \tag{5.20}$$

$$\mathbf{F}_t^{\text{prec}} = (1, 1, 0, 1, 0)' \tag{5.21}$$

$$\mathbf{G}_t^{\text{prec}} = \text{blockdiag}(1, \mathbf{G}_{\alpha,1}^{\text{prec}}, \mathbf{G}_{\alpha,2}^{\text{prec}}). \tag{5.22}$$

We specify the evolution covariance, $\mathbf{W}_t^{\text{prec}}$, using the two discount factors $\boldsymbol{\delta}^{\text{prec}} = (0.99, 0.94)$ for the baseline and seasonal components, and fix $\sigma_{\text{prec}} = 0.28$ to facilitate convergence.

Lastly, we model the 0.85-quantile of the temperature time-series again with a first order baseline component, η_t^{temp}, and a Fourier form seasonal component this time including the annual, semi-annual and tri-annual harmonics, $\alpha_{1,t}^{\text{temp}(l)}$ for $l = 1, 2, 3$, respectively. These components for the temperature quantile can be represented as

$$\boldsymbol{\theta}_t^{\text{temp}} = (\eta_t^{\text{temp}}, \alpha_{1,t}^{\text{temp}(1)}, \alpha_{2,t}^{\text{temp}(1)}, \alpha_{1,t}^{\text{temp}(2)}, \alpha_{2,t}^{\text{temp}(2)}, \alpha_{1,t}^{\text{temp}(3)}, \alpha_{2,t}^{\text{temp}(3)})' \tag{5.23}$$

$$\mathbf{F}_t^{\text{temp}} = (1, 1, 0, 1, 0, 1, 0)' \tag{5.24}$$

$$\mathbf{G}_t^{\text{temp}} = \text{blockdiag}(1, \mathbf{G}_{\alpha,1}^{\text{temp}}, \mathbf{G}_{\alpha,2}^{\text{temp}}, \mathbf{G}_{\alpha,3}^{\text{temp}}). \tag{5.25}$$

Finally, we specify the evolution covariance, $\mathbf{W}_t^{\text{temp}}$, using the two discount factors $\boldsymbol{\delta}^{\text{temp}} = (0.97, 0.92)$ for the baseline and seasonal components, and fix $\sigma_{\text{temp}} = 0.32$ to facilitate convergence.

To specify $\{\boldsymbol{\theta}_t, \mathbf{F}_t, \mathbf{G}_t, \mathbf{W}_t\}$ found in Equations (5.8)-(5.9), we combine the

individual evolutions discussed above as follows:

$$\boldsymbol{\theta}'_t = (\boldsymbol{\theta}_t^{\mathrm{BT}'}, \boldsymbol{\theta}_t^{\mathrm{prec}'}, \boldsymbol{\theta}_t^{\mathrm{temp}'}) \tag{5.26}$$

$$\mathbf{F}'_t = \mathrm{blockdiag}(\mathbf{F}_t^{\mathrm{BT}'}, \mathbf{F}_t^{\mathrm{prec}'}, \mathbf{F}_t^{\mathrm{temp}'}) \tag{5.27}$$

$$\mathbf{G}_t = \mathrm{blockdiag}(\mathbf{G}_t^{\mathrm{BT}}, \mathbf{G}_t^{\mathrm{prec}}, \mathbf{G}_t^{\mathrm{temp}}) \tag{5.28}$$

$$\mathbf{W}_t = \mathrm{blockdiag}(\mathbf{W}_t^{\mathrm{BT}}, \mathbf{W}_t^{\mathrm{prec}}, \mathbf{W}_t^{\mathrm{temp}}). \tag{5.29}$$

This results in a state evolution of dimension $q = 15$. Specifying $\mathbf{W}_t$ in this way allows for component discounting with discount factors

$$\boldsymbol{\delta} = \{0.86, 0.82, 0.99, 0.94, 0.97, 0.92\}$$

of dimensions $\{1, 2, 1, 4, 1, 6\}$, respectively. We complete the models with conjugate prior $\boldsymbol{\theta}_0 \sim \mathrm{N}(\mathbf{m}_0, \mathbf{C}_0)$ with $\mathbf{m}_0 = \mathbf{0}$ and $\mathbf{C}_0 = I_q$. Further, $\gamma_{\mathrm{BT}}, \gamma_{\mathrm{prec}} \sim t_{(-0.213,5.137)}(0, 1)$ and $\gamma_{\mathrm{temp}} \sim t_{(-5.137,0.213)}(0, 1)$, all with 1 degree of freedom. Lastly, we can combine the fixed scale parameter values into the single parameter vector, $\boldsymbol{\sigma} = (0.1, 0.28, 0.32)'$.

5.4.2 Results

We apply our MVexDQLM with the structure defined above to estimate the 0.15, 0.15 and 0.85-quantiles of the Big Tree water flow (on the log-scale), precipitation, and temperature time-series, respectively. The results in this section are computed using the MCMC Algorithm 3 and are implemented in the R programming language on a personal computer with a 2.5 GHz Intel Core i5 processor. We run each model for 1000 iterations after a burn-in of 1000 iterations. For three time-series of length 1010, the computation times are approximately two and three hours for the univariate and multivariate models, respectfully, with the

	M_0	M_1
γ_{BT}	-0.0957 (-0.0993,-0.0918)	-0.0291 (-0.0338,-0.0256)
γ_{prec}	-0.1056 (-0.1092,-0.1024)	-0.0366 (-0.0410,-0.0313)
γ_{temp}	-3.059 (-3.086,-3.034)	-3.138 (-3.158,-3.111)
$\mathbf{R}_{BT,prec}$	—	0.943 (0.930,0.959)
$\mathbf{R}_{BT,temp}$	—	-0.397 (-0.491,-0.316)
$\mathbf{R}_{prec,temp}$	—	-0.482 (-0.561,-0.419)
KL_{BT}	0.1903	0.0035
KL_{prec}	0.4922	0.0019
KL_{temp}	0.1558	0.0054
pplc	3219.24	2267.26
run-time	2.04	2.87

Table 5.2: Big Tree log-water flow, precipitation, and temperature analysis results for M_0 and M_1. Posterior summaries (format: mean (95% CrI)) for skewness parameters γ_{BT} (BT water flow), γ_{prec} (precipitation), γ_{temp} (temperature), as well as the latent correlations $\mathbf{R}_{\text{BT,prec}}$ (BT flow and precipitation), $\mathbf{R}_{\text{BT,temp}}$ (BT flow and temperature), $\mathbf{R}_{\text{prec,temp}}$ (precipitation and temperature). pplc: Posterior predictiveloss criterion under the check loss function. KL: Kullback-Liebler divergences of the marginal one-step-ahead distributions. Run-time: MCMC run-times in hours.

exact times found in Table 5.2.

Also found in Table 5.2 are the 95% CrIs of the skewness parameters $\boldsymbol{\gamma}$. The CrIs of both M_1 and M_0 are seen to be distinct from zero, as well as distinct from one another. Also seen in Table 5.2 are estimates from the latent correlation $\mathbf{R}$. In particular, we see a strong latent correlation between the BT water flow and precipitation, as expected. Also expected, we see moderate negative latent correlations between the BT water flow and temperature, as well as between precipitation and temperature.

Figure 5.4 illustrates the MAP quantiles of M_1, the multivariate model. The baseline and seasonal components of each time-series are decomposed in Figure 5.5. It can be seen that the baseline component of the BT water flow 0.15-quantile estimate varies throughout the time period, with a slight downward trend. There is a notable period of prolonged low water flow seen around 1987 to 1992 when CA

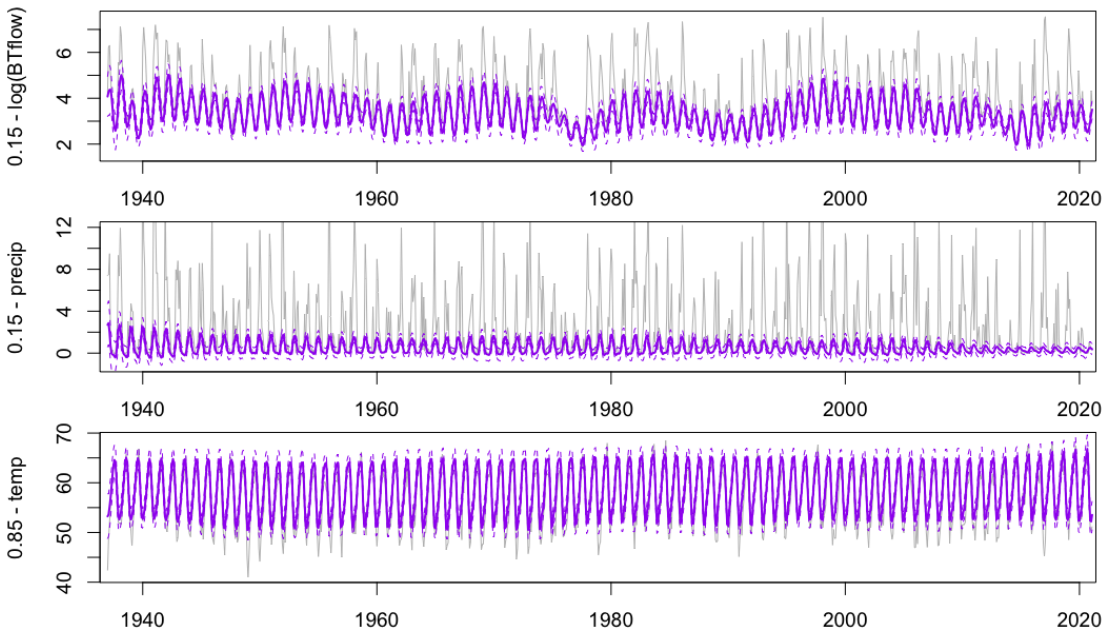

Figure 5.4: Data with overlaid solid line indicating the M_1 MAP quantiles with surrounding dashed lines indicated in the 95% CrIs. Top panel: log-Big Tree water flow data with 0.15-quantile estimates. Middle panel: Precipitation data with 0.15-quantile estimates. Bottom panel: Temperature data with 0.85-quantile estimates.

experienced a significant statewide multi-year drought (U.S. National Integrated Drought Information System, 2008). During this and other periods of low water flow, the seasonal component of the quantile estimate also exhibits a decrease in amplitude. Also seen in Figure 5.5, the precipitation 0.15-quantile baseline estimate illustrates a clear and consistent decrease. A consistent decrease in the seasonal amplitudes of the precipitation estimates is seen as well. Alternatively, the baseline estimate of the 0.85-quantile of the temperature time-series exhibits a general upward trend, with the seasonal variation remaining relatively consistent throughout the time period considered.

Figure 5.5 also includes the analogous components from the univariate model, M_0, for comparison. It can be seen that the 95% CrIs of the seasonal components overlap for all time-series, to the extent that they can be difficult to visually

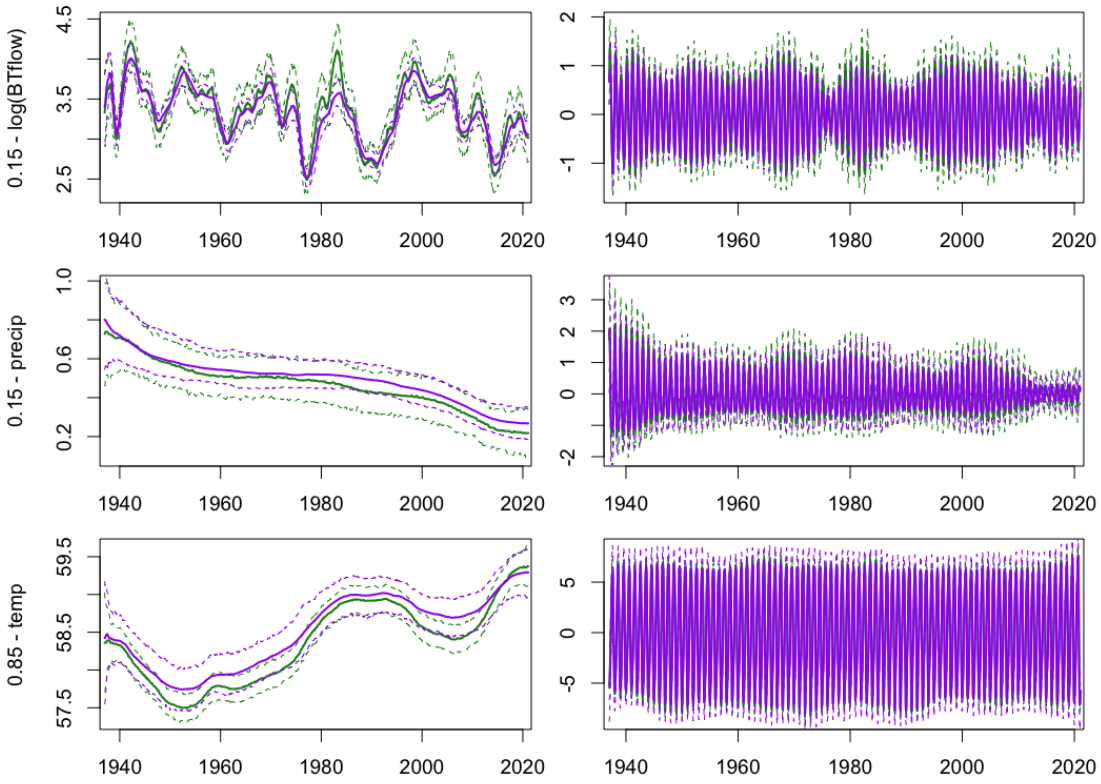

Figure 5.5: Baseline (first column) and seasonal (second column) components of the quantiles estimated by M_0 (purple) and M_0 (green). Solid lines indicate the MAP estimates with surrounding dashed lines indicated in the 95% CrIs. Top row: log-Big Tree water flow first order-polynomial baseline and seasonal component including harmonic $l = 1$. Middle row: Precipitation first order-polynomial baseline and seasonal component including harmonics $l = 1, 2$. Bottom row: Temperature first order-polynomial baseline and seasonal component including harmonics $l = 1, 2, 3$.

distinguish from one another. Further, the 95% CrIs of the baseline components of the two models also overlap for all time-series, with one exception of the BT water flow in the mid-eighties.

To assess the predictive value added by the multivariate structure in M_1, we compare the pplc and KL divergence of the marginal one-step-ahead forecast distributions for M_1 to those of M_0, seen in Table 5.2. M_1 is favored with smaller values of both comparison criteria, suggesting the multivariate structure improves

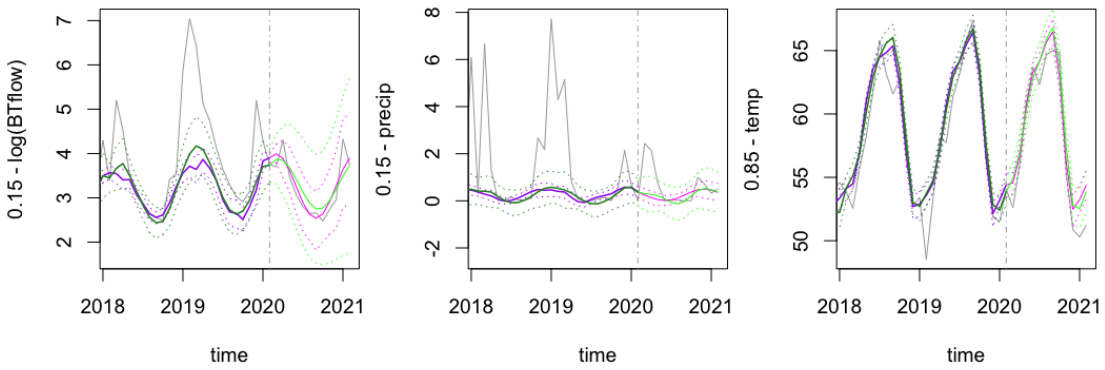

Figure 5.6: 12-step-ahead quantile forecasts for the Big Tree log-water flow, precipitation, and temperature data (left to right). Forecasts begin February 2020 through February 2021 and are overlaid on the respective time-series. The vertical dot-dashed line is at the beginning of the forecast period, February 2020, for reference. Solid lines indicate mean estimates and dotted lines 95% CrI, with the filtered estimates seen leading up to February 2020 and the forecast estimates beyond February 2020. Results from M_1 are seen in purple and magenta, and results from M_0 are seen in dark and light green.

both the predictive and forecasting power of M_1. In particular, the marginal KL divergences of M_1 are very close to zero suggesting the transformed one-step-ahead predictive distribution estimates discussed in Section 5.3 are very near normality. To further investigate the gain added by the correlation structure, we examine the k-step-ahead quantile forecast distributions. For each time t the k-step-ahead future marginal distribution of the quantiles is

$$\mathbf{F}'_{t+k}\boldsymbol{\theta}_{t+k}|\boldsymbol{\xi}_{-\boldsymbol{\theta}_{1:T}}, y_1, \ldots, y_t \sim \mathrm{N}_n(\mathbf{F}'_{t+k}\mathbf{a}_t(k), \mathbf{F}'_{t+k}\mathbf{R}_t(k)\mathbf{F}_{t+k}) \tag{5.30}$$

where $\mathbf{a}_t(k) = \mathbf{G}_{t+k}\mathbf{a}_t(k-1)$, $\mathbf{R}_t(k) = \mathbf{G}_{t+k}\mathbf{R}_t(k-1)\mathbf{G}'_{t+k} + \mathbf{W}_{t+k}$, $\mathbf{a}_t(0) = \mathbf{m}_t$, and $\mathbf{R}_t(0) = \mathbf{C}_t$, with $\mathbf{m}_t$ and $\mathbf{C}_t$ denoting the filtered mean and covariance of $\boldsymbol{\theta}_t$, respectively. Contrary to the k-step-ahead quantile forecast distributions of the exDQLM defined in Equation 3.18, here $\mathbf{F}'_{t+k}\boldsymbol{\theta}_{t+k}$ is n dimensional, thus the above distribution provides forecast estimates for all $n = 3$ quantiles. The posterior means and 95% CrIs of these distributions for 12-steps-ahead can be seen

in Figure 5.6. The MAP quantile forecasts of M_0 are almost entirely contained within the 95% CrIs of the M_1 forecasts with few exceptions. However, the 95% CrIs of M_1 are markedly smaller than those of M_0 suggesting the multivariate model provides more precise forecasts over extended periods.

5.5 Multivariate IVT analysis

In this section, we expand our analysis of the Santa Cruz average daily magnitude IVT 0.85-quantile threshold seen in Chapter 3 to 64 locations spanning all of California. The gridded observation locations used in this analysis are depicted in Figure 5.7 and correspond to the ERA-Interim IVT dataset used in Guan and Waliser (2015), discussed in Section 3.1. An example of the correlation between the time-series at nearby locations can be seen in Figure 5.8, which illustrates the average daily magnitude IVT time-series near Santa Cruz (SC), CA as well as near Fort Bragg (FB) and Santa Barbara (SB), CA. The locations of these observed time series are indicated in Figure 5.7. To incorporate this correlation structure between observation locations, we apply the MVexDQLM to jointly estimate the 0.85-quantiles of the magnitude IVT over all 64 observed locations. Again we include the immediate and lagged effects of ELI in the analysis, this time via a single transfer function component common to all locations. The ELI time series used can be seen in Figure 3.1 of Chapter 3.

5.5.1 Modeling Framework

In this application, $p_{0j} = 0.85$ for $j = 1, \ldots, n$ with $n = 64$. For all n time-series, we define the structure of the 0.85-quantile to include a second order baseline as well as seasonal components for harmonics $l = 1, 2, 4$. Thus for the jth

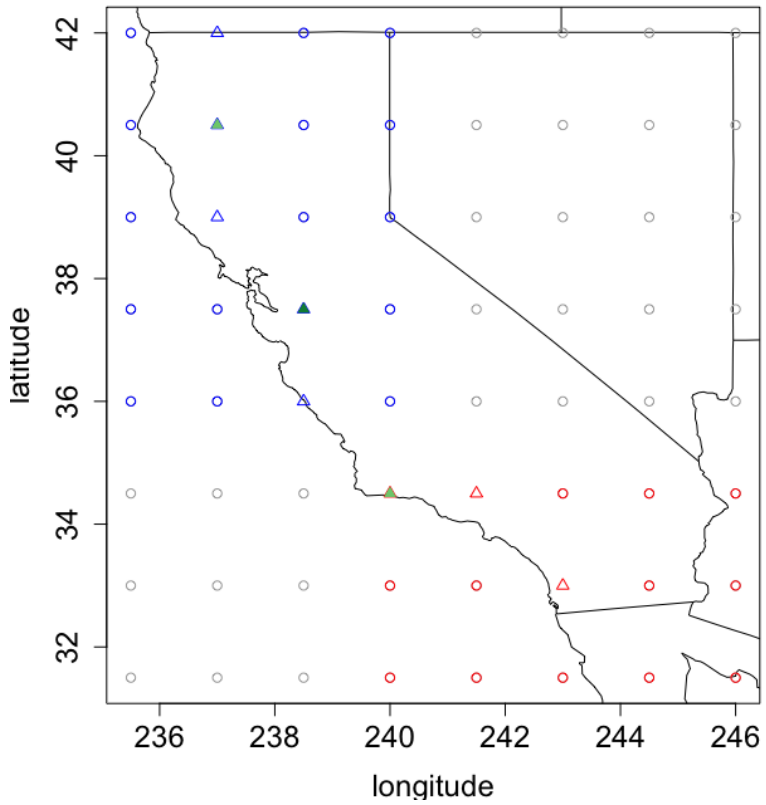

Figure 5.7: IVT magnitude observation locations spanning all of CA. Triangles indicate coastal locations at which Guan and Waliser (2015) detect ARs to make landfall. The dark green filled triangle marks the location nearest Santa Cruz CA. The light green filled triangles mark the locations nearest Fort Bragg (in Northern CA) and Santa Barbara (in Southern CA). Finally, the locations outlined in blue and red are those we label as climatologically Northern and Southern CA, respectively.

location, the $q_j = 8$ dimensional evolution of the quantile is defined by

$$\boldsymbol{\theta}_t{}^j = (\eta_{1,t}^j, \eta_{2,t}^j, \alpha_{1,t}^{j(1)}, \alpha_{2,t}^{j(1)}, \alpha_{1,t}^{j(2)}, \alpha_{2,t}^{j(2)}, \alpha_{1,t}^{j(4)}, \alpha_{2,t}^{j(4)})', \tag{5.31}$$

$$\mathbf{F}^{j\prime} = (1,0,1,0,1,0,1,0), \tag{5.32}$$

$$\mathbf{G}^j = \text{blockdiag}(\mathbf{G}^\eta, \mathbf{G}^{\alpha,1}, \mathbf{G}^{\alpha,2}, \mathbf{G}^{\alpha,4}), \tag{5.33}$$

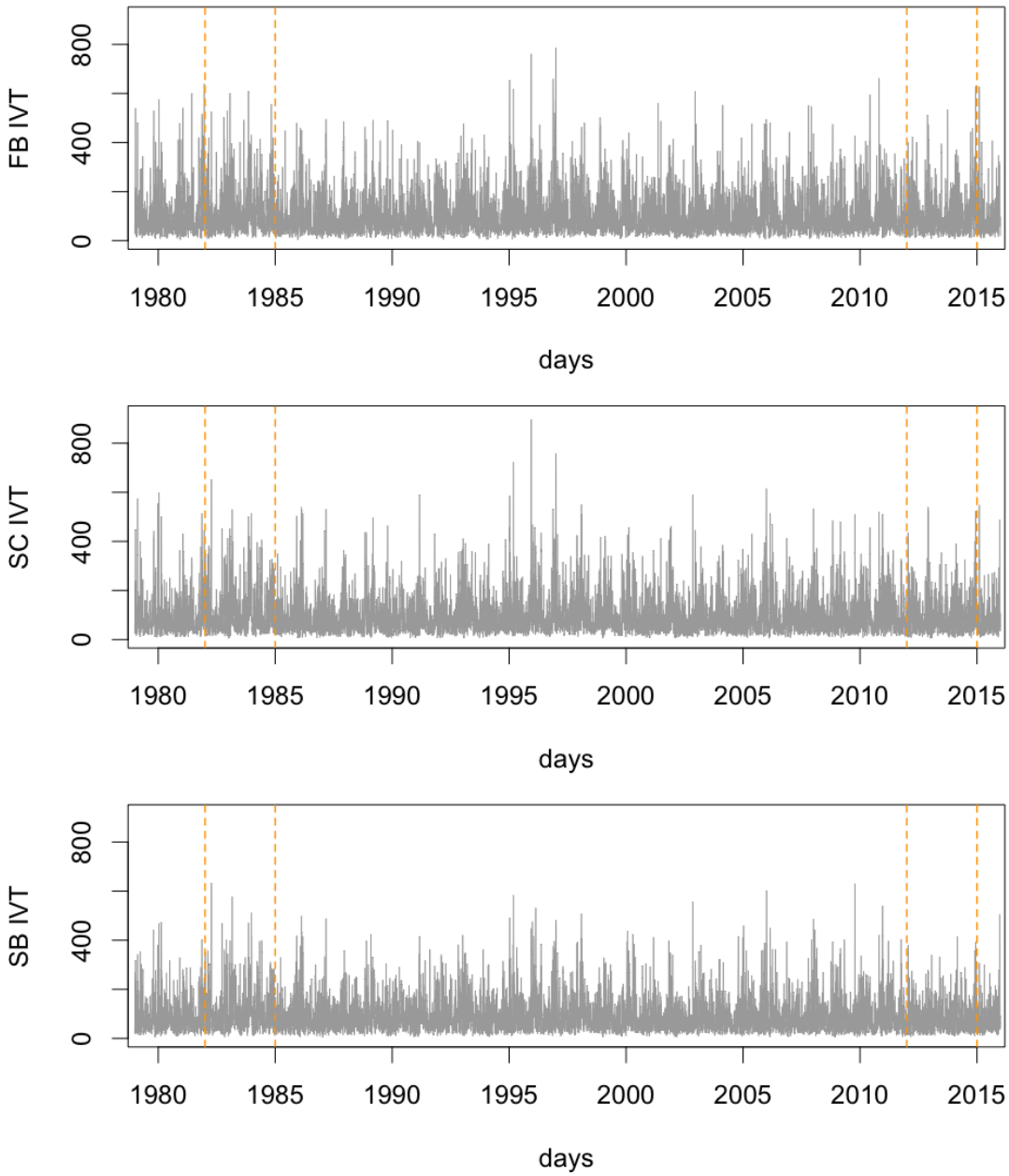

Figure 5.8: Average daily IVT magnitude near Fort Bragg CA (FB; top panel), Santa Cruz CA (SC; middle panel), and Santa Barbara CA (SB; bottom panel), spanning from 1979 through 2015. The dashed vertical lines enclose the years 1982 to 1985, a time period in which CA saw an exceptional amount of rain, and years 2012 to 2015, a period which was exceptionally dry.

where $\mathbf{G}^{\eta} = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\mathbf{G}_{\alpha,l}$ denotes the lth seasonal evolution matrix, i.e. $\mathbf{G}_{\alpha,l} = \begin{pmatrix} \cos(\frac{2\pi}{365}l) & \sin(\frac{2\pi}{365}l) \\ -\sin(\frac{2\pi}{365}l) & \cos(\frac{2\pi}{365}l) \end{pmatrix}$. Noting that $\mathbf{F}^{j}$ and $\mathbf{G}^{j}$ are not dependent on location, we can specify the baseline and seasonal evolutions of the vectorized spatial field as follows:

$$\boldsymbol{\theta}'_t = (\boldsymbol{\theta}_t^{1'}, \ldots, \boldsymbol{\theta}_t^{64'}), \tag{5.34}$$

$$\mathbf{F}' = (\mathbf{I}_n \otimes \mathbf{F}^{j'}), \tag{5.35}$$

$$\mathbf{G} = (\mathbf{I}_n \otimes \mathbf{G}^{j}), \tag{5.36}$$

$$\mathbf{W}_t = \text{blockdiag}(\mathbf{W}_t^1, \ldots, \mathbf{W}_t^{64}). \tag{5.37}$$

Here $\mathbf{I}_n$ is the $n \times n$ identity matrix, and $\otimes$ denotes the Kronecker product.

In addition to the baseline and seasonal components, we include the effects of ELI as an exponentially decaying transfer function component common to all locations. More specifically, if X_t denotes the value of ELI at time t, the transfer function MVexDQLM can be written

$$\mathbf{y}_t|\boldsymbol{\theta}_t, \boldsymbol{\gamma}, \boldsymbol{\sigma} \sim \text{MVexAL}_{p_0}(\mathbf{F}'\boldsymbol{\theta}_t + \zeta_t \mathbf{1}_n, \boldsymbol{\sigma}, \boldsymbol{\gamma}) \tag{5.38}$$

$$\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \mathbf{W}_t \sim \text{N}(\mathbf{G}\boldsymbol{\theta}_{t-1}, \mathbf{W}_t) \tag{5.39}$$

$$\zeta_t|\zeta_{t-1}, \psi_{t-1}, \omega_t \sim \text{N}(\lambda\zeta_{t-1} + X_t\psi_{t-1}, \omega_t) \tag{5.40}$$

$$\psi_t|\psi_{t-1}, \nu_t \sim \text{N}(\psi_{t-1}, \nu_t). \tag{5.41}$$

where $\boldsymbol{\theta}_t, \mathbf{F}, \mathbf{G}$, and $\mathbf{W}_t$ are as defined in Equations 5.34-5.37, and $\mathbf{1}_n$ denotes a n-dimensional vector of ones. Conditioning on optimal λ, this transfer function MVexDQLM can be rewritten in the form of a standard MVexDQLM by replacing

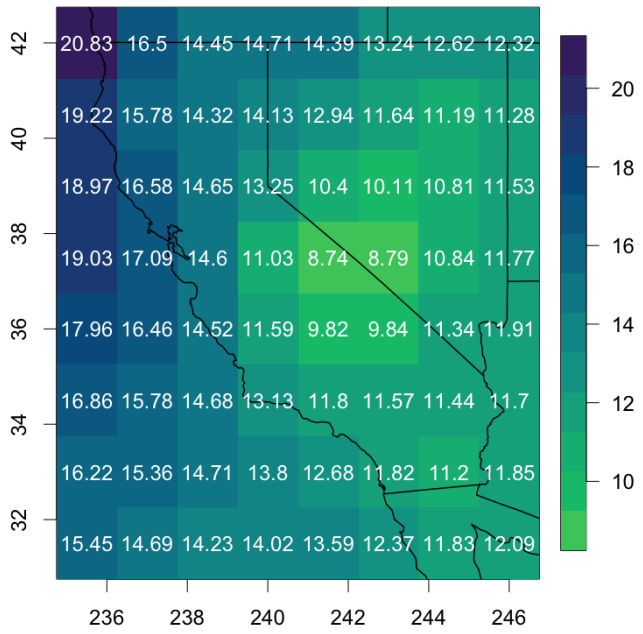

Figure 5.9: Fixed values of $\boldsymbol{\sigma}$ displayed by latitude (y-axis) and longitude (x-axis).

$\{\boldsymbol{\theta}_t, \mathbf{F}, \mathbf{G}, \mathbf{W}_t\}$ in Equations (5.8)-(5.9) with $\{\tilde{\boldsymbol{\theta}}_t, \tilde{\mathbf{F}}, \tilde{\mathbf{G}}, \tilde{\mathbf{W}}_t\}$ as follows:

$$\tilde{\mathbf{F}}'_t = (\mathbf{F}'_t, \mathbf{1}_n, \mathbf{0}_n), \tag{5.42}$$

$$\tilde{\boldsymbol{\theta}}'_t = (\boldsymbol{\theta}'_t, \zeta_t, \psi_t), \tag{5.43}$$

$$\tilde{\mathbf{G}}_t = \text{blockdiag}\left\{\mathbf{G}_t, \begin{pmatrix} \lambda & X_t \\ 0 & 1 \end{pmatrix}\right\}, \tag{5.44}$$

$$\tilde{\mathbf{W}}_t = \text{blockdiag}\left\{\mathbf{W}_t, \begin{pmatrix} \omega_t & 0 \\ 0 & \nu_t \end{pmatrix}\right\}. \tag{5.45}$$

Here $\mathbf{0}_n$ denotes a n-dimensional vector of zeros. This results in a $q = nq_j + 2 = 64(8) + 2 = 514$ dimensional state-vector, $\tilde{\boldsymbol{\theta}}'_t$. Augmenting the state-space model in this manner preserves our ability to update the state-space vector using FFBSm as well as the use of discount factors to specify the evolution covariance $\tilde{\mathbf{W}}_t$.

Similar to our approach in Chapter 2, we choose to model the baseline and seasonal components to be non-time-varying (using discount factor values of 1)

ϕ	0.486 (0.473,0.492)
$\tilde{\lambda}$	0.37
$(\delta_\zeta, \delta_\psi)$	(0.96,0.97)
run-time	25.6

Table 5.3: Multivariate IVT analysis results. Posterior summaries for exponential correlation function parameter ϕ in the format: mean (95% CrI). Optimal values of $\tilde{\lambda}, \delta_\zeta, \delta_\psi$. Run-time: ISVB run-time in hours.

such that any variability from year to year will solely be attributed the effects of ELI. The remaining discount factors for the evolution of ζ_t and ψ_t, δ_ζ and δ_ψ respectively, as well as the rate parameter λ are optimized using the sum (over j) of the one-step-ahead predictive distribution function KL divergences discussed in Section 5.3. Optimal $\lambda, \delta_\zeta, \delta_\psi$ can be found in Table 5.3.

We complete the model with conjugate priors, where applicable; $\boldsymbol{\theta}_0 \sim \text{N}(\mathbf{m}_0, \mathbf{C}_0)$ with $\mathbf{m}_0 = \mathbf{0}$ and $\mathbf{C}_0 = 100\mathbf{I}_q$, $\zeta_0 \sim N(m_{\zeta_0}, C_{\zeta_0})$ and $\psi_0 \sim N(m_{\psi_0}, C_{\psi_0})$ with $m_{\zeta_0} = m_{\psi_0} = 0$ and $C_{\zeta_0} = C_{\psi_0} = 10$. Further, $\gamma_j \sim t_{(-5.137,0.213)}(0,1)$ for all j, and σ_j is fixed at $0.75\hat{\sigma}_{j,\gamma_j=0}$ where $\hat{\sigma}_{j,\gamma_j=0}$ denotes the MAP estimate of σ_j under the DQLM, as discussed in Section 2.2.5. The fixed values of $\boldsymbol{\sigma}$ are illustrated in Figure 5.9.

5.5.2 Incorporating spatial correlation

To model the spatial dependence between locations, we apply a correlation function to define the latent correlation between observations based on the spatial distance between them. In particular, we apply the exponential correlation function with parameter ϕ. That is, if we denote the distance (in 1000 km) between the ith and jth locations by d_{ij}, then the ith and jth element of the latent

	235.5	237	238.5	240	241.5	243	244.5	246
42	(-2.22,-2.18)	(-2.18,-2.15)	(-2.28,-2.26)	(-2.33,-2.30)	(-2.36,-2.33)	(-2.43,-2.40)	(-2.55,-2.52)	(-2.58,-2.55)
40.5	(-2.29,-2.27)	(-2.16,-2.13)	(-2.22,-2.19)	(-2.23,-2.20)	(-2.30,-2.26)	(-2.41,-2.38)	(-2.53,-2.49)	(-2.56,-2.53)
39	(-2.38,-2.35)	(-2.15,-2.12)	(-2.09,-2.06)	(-2.18,-2.14)	(-2.36,-2.34)	(-2.49,-2.46)	(-2.55,-2.52)	(-2.55,-2.53)
37.5	(-2.40,-2.38)	(-2.28,-2.25)	(-2.13,-2.10)	(-2.11,-2.08)	(-2.35,-2.32)	(-2.51,-2.48)	(-2.53,-2.50)	(-2.48,-2.45)
36	(-2.45,-2.42)	(-2.42,-2.39)	(-2.30,-2.27)	(-2.20,-2.17)	(-2.28,-2.24)	(-2.42,-2.39)	(-2.48,-2.45)	(-2.42,-2.39)
34.5	(-2.49,-2.46)	(-2.50,-2.47)	(-2.43,-2.39)	(-2.31,-2.28)	(-2.21,-2.18)	(-2.33,-2.31)	(-2.46,-2.43)	(-2.44,-2.40)
33	(-2.48,-2.45)	(-2.53,-2.50)	(-2.54,-2.51)	(-2.46,-2.43)	(-2.29,-2.26)	(-2.28,-2.24)	(-2.38,-2.36)	(-2.43,-2.40)
31.5	(-2.45,-2.43)	(-2.50,-2.47)	(-2.53,-2.50)	(-2.50,-2.48)	(-2.39,-2.37)	(-2.31,-2.28)	(-2.28,-2.25)	(-2.36,-2.33)

Table 5.4: 95% CrI estimates of $\boldsymbol{\gamma}$ indexed by latitude (rows) and longitude (columns).

correlation $\mathbf{R}$ is defined by

$$R_{ij} = \exp(-d_{ij}/\phi). \tag{5.46}$$

Here, the parameter ϕ controls how fast the correlation decays with distance. More flexible families of correlation functions exist. An example of such a family is the Matérn, which is indexed by an additional parameter that controls the fine-scale decay of the correlation, i.e. how fast the correlation decreases as small distances. However, for this application we found the exponential correlation function, which is a special case of the Matérn family, produced reasonable results. We learn the parameter ϕ by replacing step 5 of the MVexDQLM ISVB algorithm with AIS Step for ϕ with proposal distribution $t^+(\langle\phi\rangle^{(k)}, \text{Var}(\phi)^{(k)})$ where t^+ denotes a Student-t distribution truncated to the positive reals with 1 degree of freedom.

5.5.3 Results

We apply our MVexDQLM ISVB algorithm to jointly estimate the 0.85-quantiles of the average daily IVT magnitude at the 64 observation locations discussed above from 1979 through 2015. The algorithm was implemented in the R programming language on a Dell PE R820 server with 32 cores (4 x Intel Xeon Sandy Bridge E5-4640 processors, each of which has 8 cores per cpu, 2.7 GHz, and 16GB RAM).

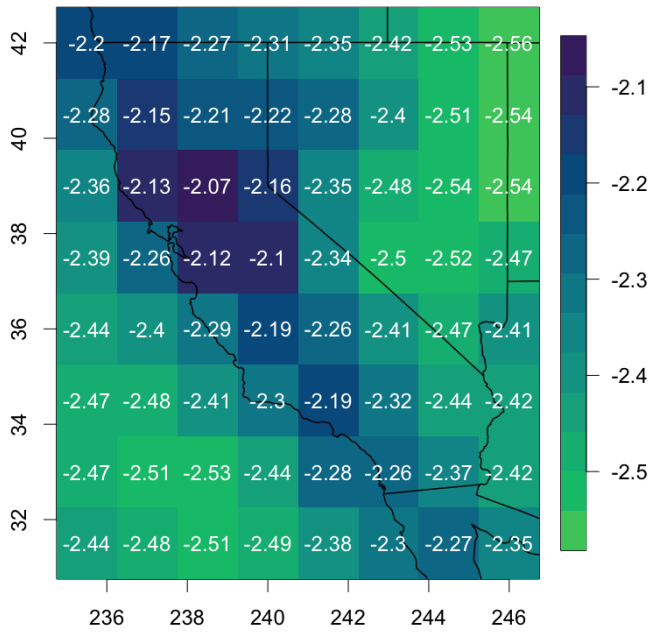

Figure 5.10: MAP estimates of $\boldsymbol{\gamma}$ displayed by latitude (y-axis) and longitude (x-axis).

For this 64-dimensional multivariate time series of length 13505, totaling 864320 observations, the ISVB computation time is approximately one day, with the exact time found in Table 5.3. Also seen in Table 5.3 are the MAP and 95% CrI estimates of the exponential correlation function parameter ϕ. The MAP estimate of 0.486 corresponds to latent correlations of approximately 0.71 at the closest observed distance (167 km; 103.8 miles) and 0.05 at the farthest observed distance (1495.3 km; 929.2 miles). Figure 5.10 depicts the MAP estimates of the skewness parameter $\boldsymbol{\gamma}$ by location. The estimates vary from -2.07 near San Francisco to -2.56 in the NE corner of Nevada. For further reference, the 95% CrIs are listed in Table 5.4, all of which are distinct from 0, thus justifying the added flexibility of the MVexAL versus the MVAL in this application.

Figure 5.11 illustrates the MAP 0.85 quantiles for the FB, SC, and SB time series, however we also examine the quantile components individually for more

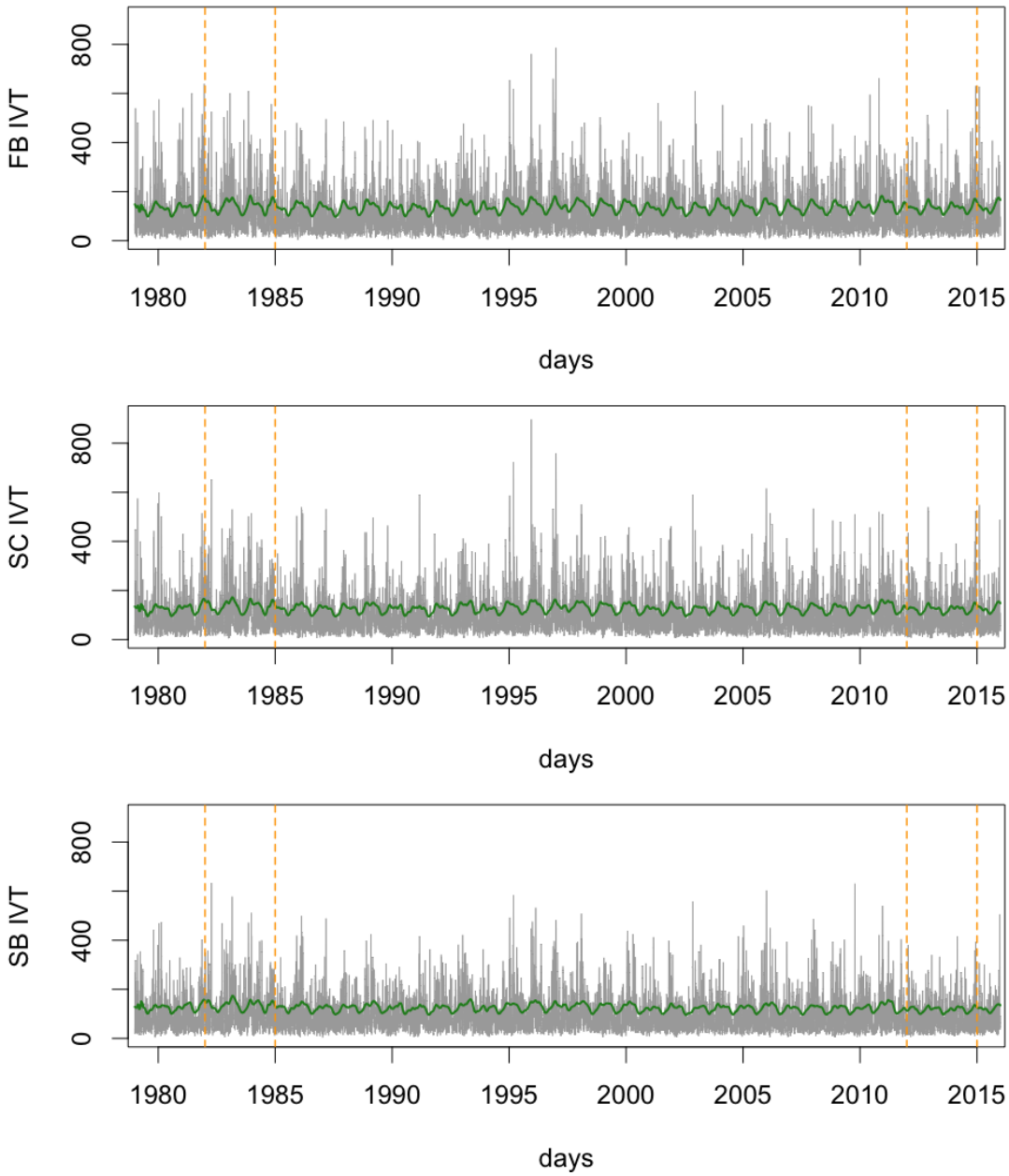

Figure 5.11: IVT data with overlaid solid line indicating the MAP 0.85-quantile estimates near Fort Bragg CA (FB; top panel), Santa Cruz CA (SC; middle panel), and Santa Barbara CA (SB; bottom panel). The dashed vertical lines enclose the years 1982 to 1985 and 2012 to 2015, time periods in which CA was exceptionally wet and dry, respectively.

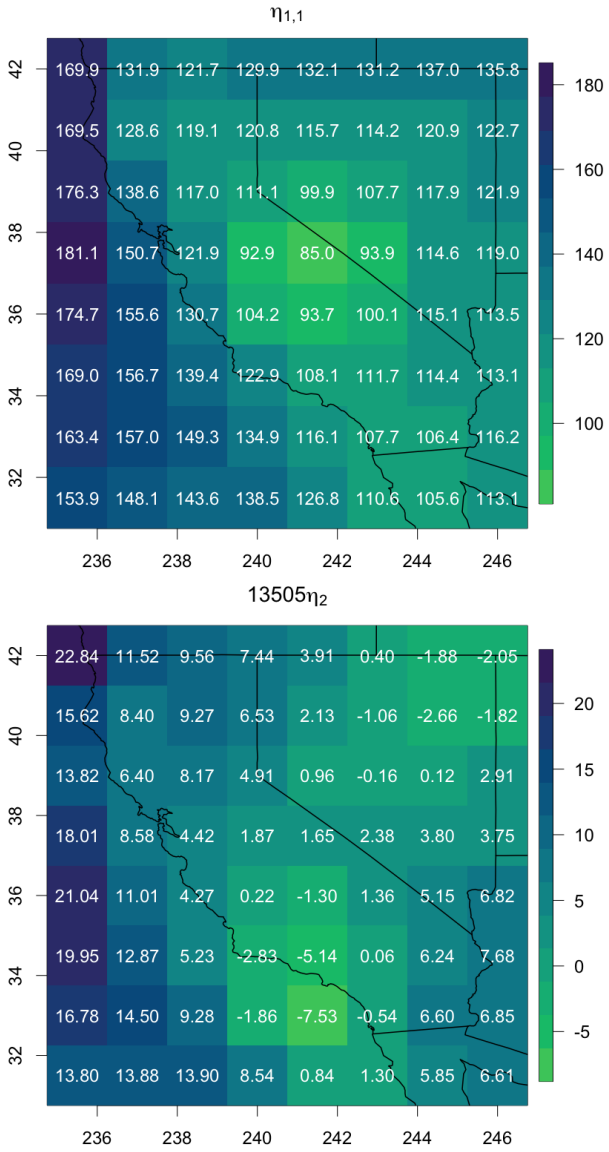

Figure 5.12: MAP estimates of the second-order polynomial trend components displayed by latitude (y-axes) and longitude (x-axes). The starting value at $t = 1$, $\eta_{1,1}$ (top), and change over the entire time period, $13505\eta_2$ (bottom).

information. In particular, MAP estimates of the baseline parameters, which capture the long-term variability of the quantile, can be found in Figure 5.12 with CrIs listed in Tables A.1 and A.2 of the Appendix A.7, for readability. Unsurprisingly, the value of the second-order polynomial trend component at location j and

time $t = 1$, $\eta^j_{1,1}$, varies dramatically from the central valley of CA to the Pacific Ocean off the coast of Northern CA, with the lowest value of 85.0 and highest of 181.1, respectively. The rate of change per time step (non-time-varying), η^j_2, is depicted in Figure 5.12 by $13505 \times \hat{\eta}^j_2$, where $\hat{\eta}^j_2$ denotes the MAP estimate. This is equivalent to the MAP long-term change in the IVT quantiles from 1979 through 2015. The change is generally positive in Northern CA, and in the Pacific Ocean. Alternatively, the change is significantly negative near Los Angeles, including several locations in the Pacific Ocean. Between the positive and negative regions, there are also locations in which the change is not significant.

From the posterior estimates of the harmonic components of the quantiles, we can compute the amplitude and phase, $A^j_l = \sqrt{\left(\alpha^{j(l)}_1\right)^2 + \left(\alpha^{j(l)}_2\right)^2}$ and $P^j_l = \arctan\left(-\alpha^{j(l)}_2/\alpha^{j(l)}_1\right)$, respectively, at each j location. MAP estimates of the amplitudes are illustrated in Figure 5.13, however tables with the amplitude CrIs, as well as tables and figures summarizing the phases be found in the Appendix A.7 for readability. The MAP estimates of the annual harmonics vary dramatically from Southern CA (0.85 in San Diego) to Northern CA (24.36 off the coast near Crescent City), suggesting the seasonal variability of the IVT is drastically different at the two extremes of CA. The semi-annual amplitudes also vary spatially, with the lowest value of 0.32 seen near Death Valley and the highest value of 15.32 off the coast north of San Francisco (SF). Lastly, the largest quarter annual amplitude of 7.37 is seen off the coast of SF and the smallest values of 1.29 is seen at the very Southern border of CA.

The amount of variability, common to all 0.85 quantiles, which can be attributed to the effects of ELI are captured with the transfer function component, ζ_t, seen in Figure 5.14. The common effects of ELI are overall significant and are dramatically more pronounced between 1982 and 1985 than between 2012 and

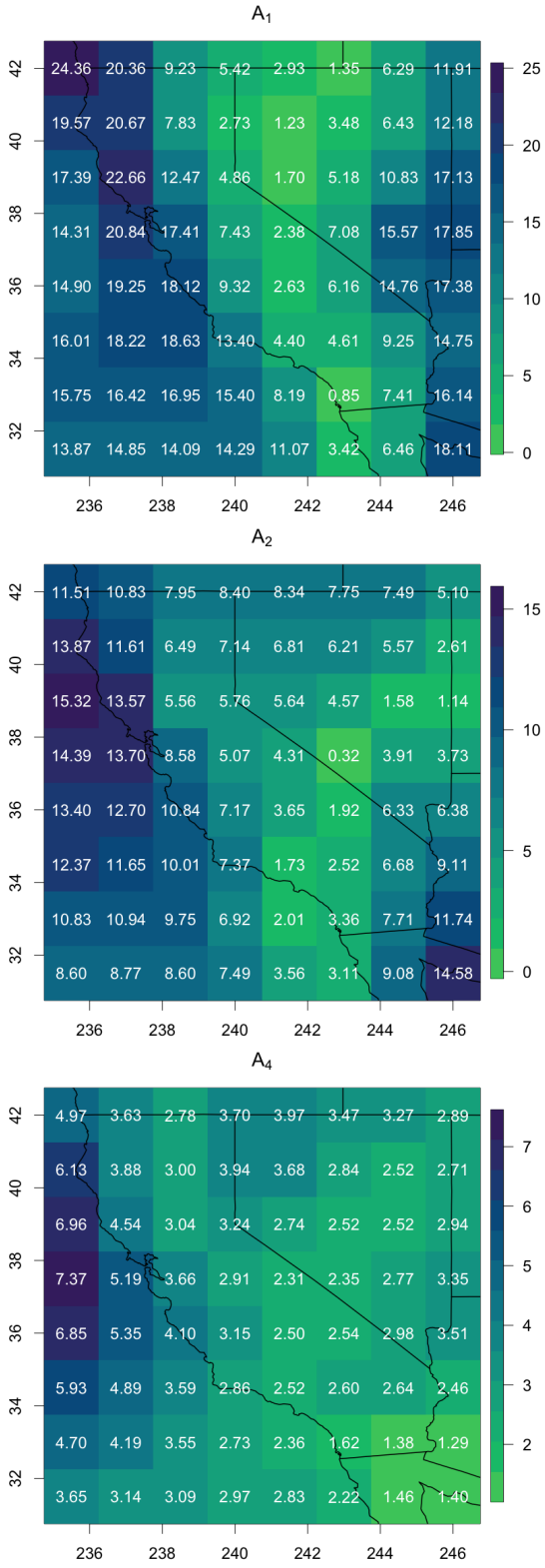

Figure 5.13: MAP estimates of the annual (top), semi-annual (middle), and quarterly (bottom) amplitudes displayed by latitude (y-axes) and longitude (x-axes).

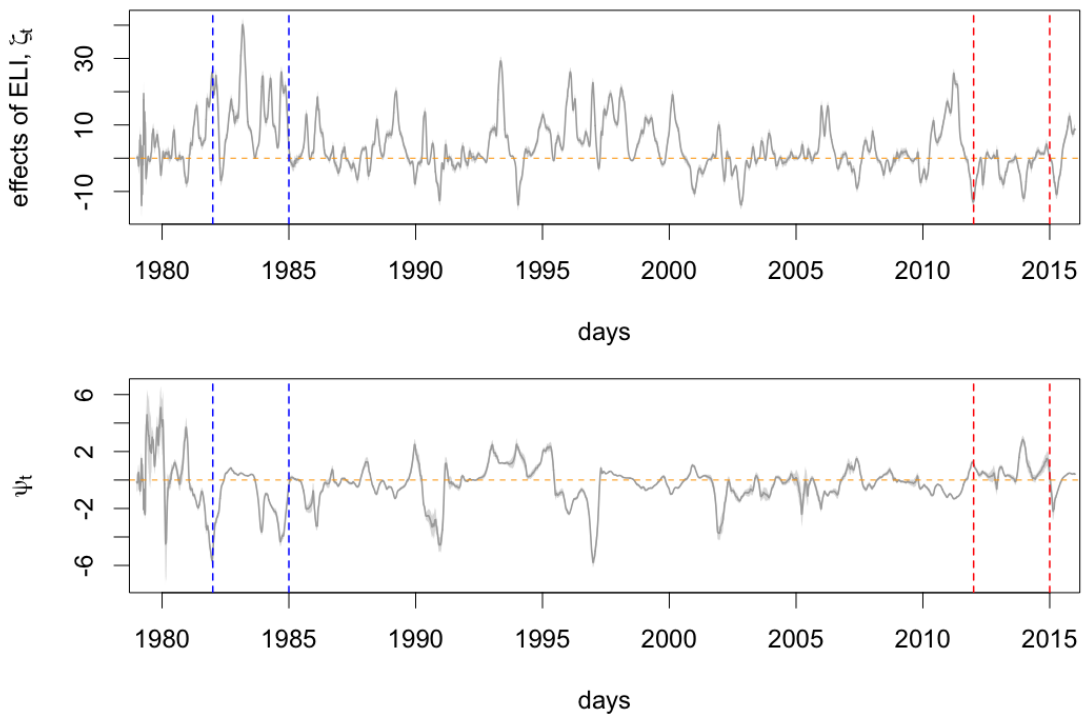

Figure 5.14: Top panel: Common effects of ELI captured by the transfer function component, ζ_t. Bottom panel: Common instantaneous effects of ELI, ψ_t. In both panels, dark grey lines indicate the MAP estimates. 95% CrI are indicated by the grey shaded regions. Dashed horizontal dashed lines are at zero, for reference. Left vertical dashed lines enclose years 1982 to 1985 in which CA saw an exceptional amount of rain. Right vertical dashed lines enclose 2012 to 2015 in which CA was exceptionally dry.

2015. Similar to the univariate analysis in Chapter 3, a majority of the effects of ELI on the IVT quantiles between 1982 and 1985 (when CA experienced heavy rainfall) are distinctly positive whereas the effects between 2012 to 2015 (when drought was widespread in CA) are negative. The instantaneous effects of ELI at time t, ψ_t, also exhibit very different behavior in the two time periods, seen in Figure 5.14. We compute the series k_t from Chapter 3 Equation (3.6) for $\epsilon = 1\text{e}{-3}$ and find the lagged effects of the common ELI component persist for around 7.5 days, on average.

Finally, with the goal of discovering any information which could contribute

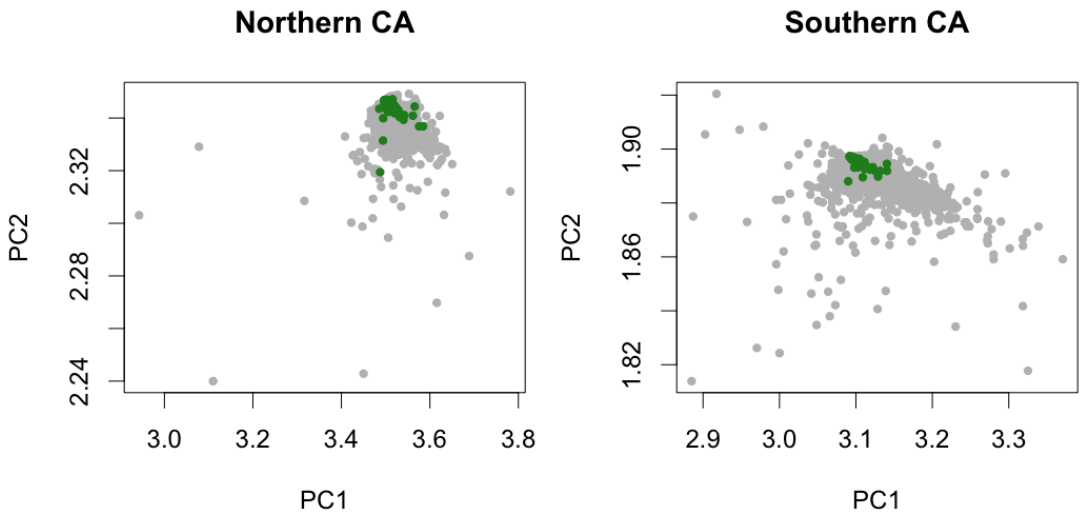

Figure 5.15: First two PCs of the components of the the forecast covariances which correspond to Northern CA (left) and Southern CA (right). Green dots indicate the PCs for times at which the considered region of CA is experiencing "high AR activity".

towards the characterization of ARs, we examine the multivariate structure further. In particular, we conduct two principle component analyses (PCA) on the forecast covariances $\mathbf{Q}_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + \tilde{\mathbf{Q}}_t$, detailed in Algorithm 4. More specifically, we consider the components of $\mathbf{Q}_t$ which correspond to locations in Northern CA (defined in Figure 5.7) and perform a PCA, then repeat for the components of $\mathbf{Q}_t$ which correspond to locations in Southern CA (also depicted in Figure 5.7). The first two principle components (PCs) for both of these analyses can be seen in Figure 5.15. We define times at which the considered region (Northern or Southern CA) experiences "high AR activity" as time steps at which at least two coastal locations within the region are indicated to have an AR make landfall by the detection scheme in Guan and Waliser (2015). The PCs from times of high AR activity exhibit a clustering behavior, as illustrated in Figure 5.15. The goal of the MVexDQLM is not classification, thus we do not pursue this further. However, it is an illustration of the rich quantitive information made possible by the

multivariate structure of the MVexDQLM.

5.6 Conclusion

We have presented several methodological and computational contributions to dynamic quantile linear modeling of multivariate time series. In particular, we developed the MVexDQLM for versatile, joint estimation of a specified quantile for each time series. Our MVexDQLM retains the advantages illustrated by the exDQLM by facilitating flexibility via the skewness parameters of the MVexAL and providing a straight-forward method for inclusion of relevant structures in the evolutions of the multivariate quantiles. The natural extension of our ISVB algorithm facilitates fast inference even for very large datasets. We illustrated the utility of our methods with the analysis of two datasets, including joint estimation of the IVT magnitude 0.85 quantile threshold at 64 locations spanning CA; an analysis made possible by our ISVB algorithm. Estimating the 0.85 quantiles with our multivariate model enhances the tools for characterization of ARs, as demonstrated by the PCA of the forecast covariances. Our methodology contributes not only to applications which require modeling of a dynamic multivariate quantile, but more generally any application with non-Guassian time-varying multivariate models.

Chapter 6

Conclusion

Within this dissertation, we have developed several new methodological and computational methods for dynamic quantile regression. All of the methods employ the exAL or its multivariate extension, the MVexAL, to promote flexibility in the estimation of the quantiles while preserving computational advantages of a parametric error distribution. We began with the development of the exDQLM for flexible estimation of a single quantile for univariate time series. We then developed the ISVB algorithm for extremely fast and accurate posterior approximation. We expanded the exDQLM to include immediate an lagged effects of a covariate via the transfer function exDQLM. The expanded model was utilized to estimate the 0.85 quantile IVT in Santa Cruz CA, while capturing both the immediate and lagged effects of ELI on the quantile. To make the these methods available to practitioners, we developed the R package, **exdqlm**, and demonstrated the utility of the software on several real datasets. Finally, we developed the MVexDQLM to consider multiple time series simultaneously and jointly estimate a specified quantile for each series. Our ISVB algorithm naturally extends to the multivariate setting, again promoting fast accurate approximate estimation. We illustrated the utility of the MVexDQLM via application to two real datasets, including an

IVT dataset spanning all of CA. A brief PCA of the multivariate structure estimated in the IVT application exhibited a clustering of the PCs for times at which certain regions of CA experienced high AR activity, thus illustrating the rich quantitative information in the multivariate structure of the MVexDQLM.

Several expansions on the work presented here could be considered. For instance, the MVexDQLM could be modified to accommodate higher dimensional problems. For example, in our multivariate analysis of the IVT dataset, vectorizing the spatial field allowed for application of our methods, including a FFBSm algorithm. However, for other spatial datasets, the number of observations may be too large, resulting in parameter vector of very high dimension, rendering the use of a standard FFBSm algorithm computationally unrealistic. A possible remedy would be the Ensemble Kalman Filter presented in Katzfuss et al. (2020), which has been successfully implemented in applications with tens of millions of dimensions. Another possible place for further development is in the selection of an adaptive importance sampling distribution for the latent parameters $\mathbf{v}_{1:T}$. Within the MVexDQLM ISVB algorithm, selecting this distribution did require a fair amount of educated guess work. More sophisticated methods for selecting the importance sampling distributions, or even estimating the variational distributions in general, could be considered. Another avenue that remains pending is the development of a spatial process based on the MVexAL. This could be done in several ways, such as extending the AL process in Lum et al. (2012). Other possible alternatives to formulate a spatial process could be considered, such as extending the approach in Tuo (2018) which considers a class of α-stable processes. A complete study comparing the properties of the new spatial process and existing methods would be of interest. Finally, it would be interesting to consider more general distributions for the quantiles. One possibility is error distributions

with tails that change in time via time-varying exAL skewness parameters. Another ambitious possibility would be the mixture distribution of exALs presented in Yan (2017) for estimation of a set of quantiles per time-series. These more general distributions would be accompanied by formidable price tags in terms of the computational tractability and model complexity, thus they would need to be coupled with further computational innovations.

Appendix A

Algorithm Details

A.1 MCMC Forward Filtering Backwards Sampling

Let $\mathbf{D}_t = \{\mathbf{y}_1, \ldots, \mathbf{y}_t\}$. To simplify the notation, we leave out conditional parameters $\mathbf{v}_{1:T}, \mathbf{s}_{1:T}, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \mathbf{R}$. For $t = 1, \ldots, T$, sample $\boldsymbol{\theta}_t|\mathbf{D}_T = \boldsymbol{\theta}_t|\mathbf{D}_T, \mathbf{v}_{1:T}, \mathbf{s}_{1:T}, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \mathbf{R}$ using the following FFBS updates:

- Forward filtering, for $t = 1, \ldots, T$ compute:
  - Prior $p(\boldsymbol{\theta}_t|\mathbf{D}_{t-1})$: Given $\boldsymbol{\theta}_{t-1}|\mathbf{D}_{t-1} \sim \mathrm{N}(\mathbf{m}_{t-1}, \mathbf{C}_{t-1})$,

$$p(\boldsymbol{\theta}_t|\mathbf{D}_{t-1}) = \int \mathrm{N}(\boldsymbol{\theta}_t|\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t)\mathrm{N}(\boldsymbol{\theta}_{t-1}|\mathbf{m}_{t-1}, \mathbf{C}_{t-1})d\boldsymbol{\theta}_{t-1}.$$

    Thus, $p(\boldsymbol{\theta}_t|\mathbf{D}_{t-1}) = \mathrm{N}(\mathbf{a}_t, \mathbf{R}_t)$, with $\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}$ and $\mathbf{R}_t = \mathbf{G}_t^T\mathbf{C}_{t-1}\mathbf{G}_t + \mathbf{W}_t$.

– Forecast $p(\mathbf{y}_t|\mathbf{D}_{t-1})$: Given $\boldsymbol{\theta}_t|D_{t-1} \sim \mathrm{N}(\mathbf{a}_t, \mathbf{R}_t)$,

$$p(\mathbf{y}_t|\mathbf{D}_{t-1}) =$$
$$\int \mathrm{N}(\mathbf{y}_t|\mathbf{F}_t'\boldsymbol{\theta}_t + \mathbf{C} \odot \boldsymbol{\sigma} \odot |\boldsymbol{\gamma}| \odot \mathbf{s}_t + \mathbf{A} \odot \mathbf{v}_t, \sqrt{\mathbf{B} \odot \mathbf{v}_t \odot \boldsymbol{\sigma}}'\mathbf{R}\sqrt{\mathbf{B} \odot \mathbf{v}_t \odot \boldsymbol{\sigma}})$$
$$\mathrm{N}(\boldsymbol{\theta}_t|\mathbf{a}_t, \mathbf{R}_t)d\boldsymbol{\theta}_t.$$

Thus, for the MVexDQLM algorithm, $p(\mathbf{y}_t|\mathbf{D}_{t-1}) = \mathrm{N}(\mathbf{f}_t, \mathbf{Q}_t)$, with

$$\mathbf{f}_t = \mathbf{F}_t'\mathbf{a}_t + \mathbf{C} \odot \boldsymbol{\sigma} \odot |\boldsymbol{\gamma}| \odot \mathbf{s}_t + \mathbf{A} \odot \mathbf{v}_t,$$

and

$$\mathbf{Q}_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + \sqrt{\mathbf{B} \odot \mathbf{v}_t \odot \boldsymbol{\sigma}}'\mathbf{R}\sqrt{\mathbf{B} \odot \mathbf{v}_t \odot \boldsymbol{\sigma}}.$$

If the data are univariate, i.e. $\mathbf{y}_t = y_t$, the forecast distribution simplifies to that used in the exDQLM algorithm: $p(y_t|D_{t-1}) = \mathrm{N}(f_t, Q_t)$, with $f_t = \mathbf{F}_t'\mathbf{a}_t + C(p)\sigma|\gamma|s_t + A(p)v_t$ and $Q_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + \sigma B(p)v_t$.

– Posterior $p(\boldsymbol{\theta}_t|\mathbf{D}_t)$: Given the prior and forecast distributions, the joint distribution can be written

$$\left[\begin{smallmatrix}\boldsymbol{\theta}_t \\ \mathbf{y}_t\end{smallmatrix}\middle|\mathbf{D}_{t-1}\right] \sim \mathrm{N}\left(\left[\begin{smallmatrix}\mathbf{a}_t \\ \mathbf{f}_t\end{smallmatrix}\right], \left[\begin{smallmatrix}\mathbf{R}_t & \mathbf{R}_t\mathbf{F}_t \\ \mathbf{F}_t'\mathbf{R}_t & \mathbf{Q}_t\end{smallmatrix}\right]\right).$$

The conditional distribution of the multivariate normal $p(\boldsymbol{\theta}_t|\mathbf{D}_t) = \mathrm{N}(\mathbf{m}_t, \mathbf{C}_t)$ with $\mathbf{m}_t = \mathbf{a}_t + \mathbf{R}_t\mathbf{F}_t\mathbf{Q}_t^{-1}(\mathbf{y}_t - \mathbf{f}_t)$ and $\mathbf{C}_t = \mathbf{R}_t - \mathbf{R}_t\mathbf{F}_t\mathbf{Q}_t^{-1}\mathbf{F}_t'\mathbf{R}_t$

- Backwards sampling, $p(\boldsymbol{\theta}_t|\mathbf{D}_T)$:

  – For T, sample $\boldsymbol{\theta}_T|\mathbf{D}_T \sim \mathrm{N}(\mathbf{m}_T, \mathbf{C}_T)$.

  – For $t = T-1, \ldots, 1$, sample $\boldsymbol{\theta}_t|\mathbf{D}_T \sim \mathrm{N}(\mathbf{m}_t^s, \mathbf{C}_t^s)$ with $\mathbf{m}_t^s = \mathbf{m}_t + \mathbf{C}_t\mathbf{G}_t'\mathbf{R}_{t+1}^{-1}(\boldsymbol{\theta}_{t+1} - \mathbf{a}_{t+1})$ and $\mathbf{C}_t^s = \mathbf{C}_t - \mathbf{C}_t\mathbf{G}_t'\mathbf{R}_{t+1}^{-1}\mathbf{G}_t\mathbf{C}_t$.

A.2 ISVB Forward Filtering Backwards Smoothing

Similarly, we can update $r^{(k+1)}(\boldsymbol{\theta}_t) = \text{N}(\mathbf{m}_t^s, \mathbf{C}_t^s)$ using FFBSm as follows.

- Forward filter, for $t = 1, \ldots, T$:
  - Prior: $r^{(k+1)}(\boldsymbol{\theta}_t|\mathbf{D}_{t-1}) = \text{N}(\mathbf{a}_t, \mathbf{R}_t)$, with $\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}$ and $\mathbf{R}_t = \mathbf{G}_t^T\mathbf{C}_{t-1}\mathbf{G}_t + \mathbf{W}_t$.
  - Forecast: In the MVexDQLM algorithm $r^{(k+1)}(\mathbf{y}_t|\mathbf{D}_{t-1}) = \text{N}(\mathbf{f}_t, \mathbf{Q}_t)$ where $\mathbf{f}_t = \mathbf{F}_t'\mathbf{a}_t + \tilde{\mathbf{Q}}_t\tilde{\mathbf{f}}_t$ and $\mathbf{Q}_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + \tilde{\mathbf{Q}}_t$. Here, the i-th element of $\tilde{\mathbf{f}}$ is

$$\tilde{\mathbf{f}}_i = \left\langle \frac{C_i|\gamma_i|}{B_i} \right\rangle^{(k)} \left\langle \frac{1}{v_{ti}} \right\rangle^{(k)} \langle s_{ti} \rangle^{(k)} \langle R^{-1} \rangle_{ii}^{(k)} + \sum_{j=1, j\neq i}^{n} \left\langle \frac{C_j|\gamma_j|\sqrt{\sigma_j}}{\sqrt{B_j B_i \sigma_i}} \right\rangle^{(k)} \left\langle \frac{1}{\sqrt{v_{tj} v_{ti}}} \right\rangle^{(k)} \langle s_{tj} \rangle^{(k)} \langle R^{-1} \rangle_{ij}^{(k)}$$

and the ij-th element of $\tilde{\mathbf{Q}}_t^{-1}$ for $i \neq j$ is

$$(\tilde{\mathbf{Q}}_t^{-1})_{ij} = \left\langle \frac{1}{\sqrt{\sigma_i \sigma_j B_i B_j}} \right\rangle^{(k)} \left\langle \frac{1}{\sqrt{v_{ti} v_{tj}}} \right\rangle^{(k)} \langle R^{-1} \rangle_{ij}^{(k)}$$

and for $i = j$

$$(\tilde{\mathbf{Q}}_t^{-1})_{ii} = \left\langle \frac{1}{\sigma_i B_i} \right\rangle^{(k)} \left\langle \frac{1}{v_{ti}} \right\rangle^{(k)} \langle R^{-1} \rangle_{ii}^{(k)}.$$

If the data are univariate, i.e. $\mathbf{y}_t = y_t$, the variational forecast distribution simplifies to that used in the exDQLM algorithm: $r^{(k+1)}(y_t|D_{t-1}) =$

$\mathrm{N}(f_t, Q_t)$, with

$$f_t = \frac{\mathbf{F}_t'\mathbf{a}_t + \left[\left\langle \frac{C(p)|\gamma|}{B(p)} \right\rangle^{(k)} \langle s_t \rangle^{(k+1)} + \left\langle \frac{A(p)}{\sigma B(p)} \right\rangle^{(k)} \Big/ \left\langle \frac{1}{v_t} \right\rangle^{(k+1)}\right]}{\left\langle \frac{1}{\sigma B(p)} \right\rangle^{(k)}}$$

$$Q_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + \left[\left\langle \frac{1}{v_t} \right\rangle^{(k+1)} \left\langle \frac{1}{\sigma B(p)} \right\rangle^{(k)}\right]^{-1}.$$

– Posterior: $r^{(k+1)}(\boldsymbol{\theta}_t|\mathbf{D}_t) = \mathrm{N}(\mathbf{m}_t, \mathbf{C}_t)$, with $\mathbf{m}_t = \mathbf{a}_t + \mathbf{R}_t\mathbf{F}_t\mathbf{Q}_t^{-1}(\mathbf{y}_t - \mathbf{f}_t)$ and $\mathbf{C}_t = \mathbf{R}_t - \mathbf{R}_t\mathbf{F}_t\mathbf{Q}_t^{-1}\mathbf{F}_t'\mathbf{R}_t$

- Backward smoother:

  – For T, $r^{(k+1)}(\boldsymbol{\theta}_T|\mathbf{D}_T) = \mathrm{N}(\mathbf{m}_T^s = \mathbf{m}_T, \mathbf{C}_T^s = \mathbf{C}_T)$

  – For $t = T-1, \ldots, 1$, $r^{(k+1)}(\boldsymbol{\theta}_t|\mathbf{D}_T) = \mathrm{N}(\mathbf{m}_t^s, \mathbf{C}_t^s)$ with $\mathbf{m}_t^s = \mathbf{m}_t + \mathbf{B}_t(\mathbf{m}_{t+1} - \mathbf{a}_{t+1})$ and $\mathbf{C}_t^s = \mathbf{C}_t + \mathbf{B}_t(\mathbf{C}_{t+1}^s - \mathbf{R}_{t+1})\mathbf{B}_t'$, where $\mathbf{B}_t = \mathbf{C}_t\mathbf{G}_t'\mathbf{R}_{t+1}^{-1}$

A.3 exDQLM ISVB Importance Sampling

The variational distribution $r^{(k+1)}(\sigma,\gamma)$ can be computed up to a proportionality constant,

$$\begin{aligned} r^{(k+1)}(\sigma,\gamma) \propto & f_0(\gamma)f_0(\sigma)\sigma^{1.5T}\exp\Big\{-\sum\langle u_t\rangle^{(k+1)}/\sigma \\ & -\frac{1}{2}\sum\Big[\frac{1}{\sigma B(p)}\left\langle\frac{1}{v_t}\right\rangle^{(k+1)}(y_t^2-2y_t\langle\mathbf{F}_t'\boldsymbol{\theta}_t\rangle^{(k+1)}+\langle(\mathbf{F}_t'\boldsymbol{\theta}_t)^2\rangle^{(k+1)}) \\ & -2(\frac{C(p)|\gamma|}{B(p)}\left\langle\frac{1}{v_t}\right\rangle^{(k+1)}\langle s_t\rangle^{(k+1)}+\frac{A(p)}{\sigma B(p)})(y_t-\langle\mathbf{F}_t'\boldsymbol{\theta}_t\rangle^{(k+1)}) \\ & +2\frac{C(p)|\gamma|A(p)}{B(p)}\langle s_t\rangle^{(k+1)}+\frac{C(p)^2\sigma|\gamma|^2}{B(p)}\langle s_t^2\rangle^{(k+1)}\left\langle\frac{1}{v_t}\right\rangle^{(k+1)} \\ & +\frac{A(p)^2}{\sigma B(p)}\langle v_t\rangle^{(k+1)}\Big]\Big\} \end{aligned}$$

where $f_0(\sigma)$ and $f_0(\gamma)$ denote the prior distributions of σ and γ, respectively. Therefore, we can update $r^{(k+1)}(\sigma,\gamma)$ with importance sampling as follows:

- For n in $1,...,N$, sample $(\sigma^n,\gamma^n)\sim l(\sigma,\gamma)$ where $l(\sigma,\gamma)$ denotes the chosen proposal distribution.
- Compute the weights

$$w(\sigma^n,\gamma^n)=\frac{r^{(k+1)}(\sigma^n,\gamma^n)}{l(\sigma^n,\gamma^n)} \tag{A.1}$$

The variational distribution $r^{(k+1)}(\sigma,\gamma)$ can be approximated by

$$r^{(k+1)}(\sigma,\gamma)\approx\frac{\sum_{n=1}^N w(\sigma^n,\gamma^n)\delta_{(\sigma^n,\gamma^n)}(\sigma,\gamma)}{\sum_{n=1}^N w(\sigma^n,\gamma^n)}. \tag{A.2}$$

Similarly, for any function $h(\sigma, \gamma)$,

$$\mathbb{E}[h(\sigma,\gamma)] \approx \frac{\sum_{n=1}^{N} h(\sigma^n,\gamma^n)w(\sigma^n,\gamma^n)}{\sum_{n=1}^{N} w(\sigma^n,\gamma^n)}. \tag{A.3}$$

A.4 exDQLM ISVB Closed Form Integrals

For notational simplicity, we omit the superscript indicating the VB iteration.

If $r(s_t) = \mathrm{N}^+(\mu_{s_t}, \sigma^2_{s_t})$, $\phi(\cdot)$ is the probability density function of the standard normal distribution, and $\Phi(\cdot)$ is its cumulative distribution function, then

$$\begin{aligned}
\langle s_t \rangle &= \mu_{s_t} + \sigma_{s_t} \frac{\phi(\mu_{s_t}/\sigma_{s_t})}{\Phi(\mu_{s_t}/\sigma_{s_t})} \\
\langle s_t^2 \rangle &= \mu_{s_t}^2 + \sigma_{s_t}^2 + \mu_{s_t}\sigma_{s_t} \frac{\phi(\mu_{s_t}/\sigma_{s_t})}{\Phi(\mu_{s_t}/\sigma_{s_t})}
\end{aligned}$$

If $r(v_t) = \mathrm{GIG}(\lambda_{v_t}, \chi_{v_t}, \psi_{v_t})$ and $K_\lambda(\cdot)$ is a modified Bessel Function of the second kind with order λ, then

$$\begin{aligned}
\langle v_t \rangle &= \frac{\sqrt{\chi_{v_t}} K_{\lambda_{v_t}+1}(\sqrt{\chi_{v_t}\psi_{v_t}})}{\sqrt{\psi_{v_t}} K_{\lambda_{v_t}}(\sqrt{\chi_{v_t}\psi_{v_t}})} \\
\left\langle \frac{1}{v_t} \right\rangle &= \frac{\sqrt{\psi_{v_t}} K_{\lambda_{v_t}+1}(\sqrt{\chi_{v_t}\psi_{v_t}})}{\sqrt{\chi_{v_t}} K_{\lambda_{v_t}}(\sqrt{\chi_{v_t}\psi_{v_t}})} - \frac{2\lambda_{v_t}}{\chi_{v_t}}
\end{aligned}$$

If $r(\boldsymbol{\theta}_t) = \mathrm{N}(\mathbf{m}_t^s, \mathbf{C}_t^s)$, then

$$\begin{aligned}
\langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle &= \mathbf{F}_t' \mathbf{m}_t^s \\
\langle (\mathbf{F}_t' \boldsymbol{\theta}_t)^2 \rangle &= \mathbf{F}_t' \mathbf{C}_t^s \mathbf{F}_t + (\mathbf{F}_t' \mathbf{m}_t^s)^2
\end{aligned}$$

Lastly, if $r(\sigma, \gamma)$ is approximated with IS according to Equation (A.2), the

following expectations can be approximated using Equation (A.3);

$$\left\langle \frac{C(p)^2\sigma|\gamma|^2}{B(p)} \right\rangle, \left\langle \frac{1}{\sigma} \right\rangle, \left\langle \frac{C(p)|\gamma|}{B(p)} \right\rangle, \left\langle \frac{C(p)|\gamma|A(p)}{B(p)} \right\rangle, \left\langle \frac{1}{\sigma B(p)} \right\rangle, \left\langle \frac{A(p)}{\sigma B(p)} \right\rangle, \text{ and } \left\langle \frac{A(p)^2}{\sigma B(p)} \right\rangle.$$

A.5 MVexDQLM ISVB Importance Sampling

The variational distributions $r(\mathbf{v}_t)^{(k+1)}$ for $t = 1, \ldots, T$, $r(\boldsymbol{\sigma}, \boldsymbol{\gamma})^{(k+1)}$, and $r(\mathbf{R})^{(k+1)}$ can be computed up to proportionality constants. Following the notation defined Section 5.2.2 and omitting the superscripts indicating the VB iteration

for notational simplicity, these distribution can be written as follows.

$$
\begin{aligned}
r(\mathbf{v}_t) \propto \prod_{i=1}^{n} \Big[v_{it}^{-1/2} \exp\Big\{ -\Big\langle \sigma_i^{-1} \Big\rangle v_{it} \Big\} \Big] \\
\exp\Big\{ -\frac{1}{2} \sum_{i=1}^{n} \Big(\frac{1}{v_{ti}} (y_{ti}^2 - 2y_{ti} \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_i + \langle (\mathbf{F}_t' \boldsymbol{\theta}_t)^2 \rangle_i) \Big\langle \frac{1}{B_i \sigma_i} \Big\rangle \langle R^{-1} \rangle_{ii} \\
+ (y_{ti} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_i) \sum_{j=1, j \neq i}^{n} \frac{1}{\sqrt{v_{ti} v_{tj}}} (y_{tj} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_j) \Big\langle \frac{1}{\sqrt{B_i \sigma_i B_j \sigma_j}} \Big\rangle \langle R^{-1} \rangle_{ij} \Big) \\
+ \sum_{i=1}^{n} \Big(\frac{1}{v_{ti}} \Big\langle \frac{C_i |\gamma_i|}{B_i} \Big\rangle \langle s_{ti} \rangle (y_{ti} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_i) \langle R^{-1} \rangle_{ii} \\
+ \langle s_{ti} \rangle \sum_{j=1, j \neq i}^{n} \frac{1}{\sqrt{v_{ti} v_{tj}}} \Big\langle \frac{C_i |\gamma_i| \sqrt{\sigma_i}}{\sqrt{B_i B_j \sigma_j}} \Big\rangle (y_{tj} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_j) \langle R^{-1} \rangle_{ij} \Big) \\
- \frac{1}{2} \sum_{i=1}^{n} \Big(\frac{1}{v_{ti}} \Big\langle \frac{C_i^2 |\gamma_i|^2 \sigma_i}{B_i} \Big\rangle \langle s_{ti}^2 \rangle \langle R^{-1} \rangle_{ii} \\
+ \sum_{j=1, j \neq i}^{n} \frac{1}{\sqrt{v_{ti} v_{tj}}} \Big\langle \frac{C_i C_j |\gamma_i \gamma_j| \sqrt{\sigma_i \sigma_j}}{\sqrt{B_i B_j}} \Big\rangle \langle s_{ti} s_{tj} \rangle \langle R^{-1} \rangle_{ij} \Big) \\
+ \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \frac{\sqrt{v_{ti}}}{\sqrt{v_{tj}}} \Big\langle \frac{A_i}{\sqrt{B_i \sigma_i B_j \sigma_j}} \Big\rangle (y_{tj} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_j) \langle R^{-1} \rangle_{ij} \\
- \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \frac{\sqrt{v_{ti}}}{\sqrt{v_{tj}}} \Big\langle \frac{A_i C_j |\gamma_j| \sqrt{\sigma_j}}{\sqrt{B_i B_j \sigma_i}} \Big\rangle \langle s_{tj} \rangle \langle R^{-1} \rangle_{ij} \\
- \frac{1}{2} \sum_{i=1}^{n} \Big(v_{ti} \Big\langle \frac{A_i^2}{B_i \sigma_i} \Big\rangle \langle R^{-1} \rangle_{ii} + \sum_{j=1, j \neq i}^{n} \sqrt{v_{ti} v_{tj}} \Big\langle \frac{A_i A_j}{\sqrt{B_i \sigma_i B_j \sigma_j}} \Big\rangle \langle R^{-1} \rangle_{ij} \Big) \Big\}.
\end{aligned}
$$

$$
\begin{aligned}
r(\boldsymbol{\sigma},\boldsymbol{\gamma}) \propto f_0(\boldsymbol{\gamma}) f_0(\boldsymbol{\sigma}) \prod_{i=1}^{n} \Big[\sigma_i^{-3T/2} \exp\Big\{ -\sum_{t=1}^{T} \langle v_{ti} \rangle / \sigma_i \Big\} \Big] \\
\exp\Big\{ -\frac{1}{2} \sum_{i=1}^{n} \Big(\frac{1}{B_i \sigma_i} (y_{ti}^2 - 2 y_{ti} \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_i + \langle (\mathbf{F}_t' \boldsymbol{\theta}_t)^2 \rangle_i) \Big\langle \frac{1}{v_{ti}} \Big\rangle \langle R^{-1} \rangle_{ii} \\
+ (y_{ti} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_i) \sum_{j=1, j \neq i}^{n} \frac{1}{\sqrt{B_i \sigma_i B_j \sigma_j}} (y_{tj} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_j) \Big\langle \frac{1}{\sqrt{v_{ti} v_{tj}}} \Big\rangle \langle R^{-1} \rangle_{ij} \Big) \\
+ \sum_{i=1}^{n} \Big(\frac{C_i |\gamma_i|}{B_i} \Big\langle \frac{1}{v_{ti}} \Big\rangle \langle s_{ti} \rangle (y_{ti} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_i) \langle R^{-1} \rangle_{ii} \\
+ \langle s_{ti} \rangle \sum_{j=1, j \neq i}^{n} \frac{C_i |\gamma_i| \sqrt{\sigma_i}}{\sqrt{B_i B_j \sigma_j}} (y_{tj} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_j) \Big\langle \frac{1}{\sqrt{v_{ti} v_{tj}}} \Big\rangle \langle R^{-1} \rangle_{ij} \Big) \\
- \frac{1}{2} \sum_{i=1}^{n} \Big(\frac{C_i^2 |\gamma_i|^2 \sigma_i}{B_i} \langle s_{ti}^2 \rangle \Big\langle \frac{1}{v_{ti}} \Big\rangle \langle R^{-1} \rangle_{ii} \\
+ \sum_{j=1, j \neq i}^{n} \frac{C_i C_j |\gamma_i \gamma_j| \sqrt{\sigma_i \sigma_j}}{\sqrt{B_i B_j}} \langle s_{ti} s_{tj} \rangle \Big\langle \frac{1}{\sqrt{v_{ti} v_{tj}}} \Big\rangle \langle R^{-1} \rangle_{ij} \Big) \\
+ \sum_{i=1}^{n} \Big(\frac{A_i}{B_i \sigma_i} (y_{ti} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_i) \langle R^{-1} \rangle_{ii} \\
+ \sum_{j=1, j \neq i}^{n} \frac{A_i}{\sqrt{B_i \sigma_i B_j \sigma_j}} \Big\langle \frac{\sqrt{v_{ti}}}{\sqrt{v_{tj}}} \Big\rangle (y_{tj} - \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle_j) \langle R^{-1} \rangle_{ij} \Big) \\
- \sum_{i=1}^{n} \Big(\frac{A_i C_i |\gamma_i|}{B_i} \langle s_{ti} \rangle \langle R^{-1} \rangle_{ii} \\
+ \sum_{j=1, j \neq i}^{n} \frac{A_i C_j |\gamma_j| \sqrt{\sigma_j}}{\sqrt{B_i B_j \sigma_i}} \langle s_{tj} \rangle \Big\langle \frac{\sqrt{v_{ti}}}{\sqrt{v_{tj}}} \Big\rangle \langle R^{-1} \rangle_{ij} \Big) \\
- \frac{1}{2} \sum_{i=1}^{n} \Big(\frac{A_i^2}{B_i \sigma_i} \langle v_{ti} \rangle \langle R^{-1} \rangle_{ii} \\
+ \sum_{j=1, j \neq i}^{n} \frac{A_i A_j}{\sqrt{B_i \sigma_i B_j \sigma_j}} \langle \sqrt{v_{ti} v_{tj}} \rangle \langle R^{-1} \rangle_{ij} \Big) \Big\}.
\end{aligned}
$$

$$
\begin{aligned}
r(\mathbf{R}) \propto\; & f_0(\mathbf{R}) \det|\mathbf{R}|^{-T/2} \\
& \exp\Bigg\{ -\frac{1}{2}\sum_{i=1}^{n}\Bigg((y_{ti}^2 - 2y_{ti}\langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_i + \langle(\mathbf{F}_t'\boldsymbol{\theta}_t)^2\rangle_i)\left\langle \frac{1}{B_i\sigma_i}\right\rangle\left\langle\frac{1}{v_{ti}}\right\rangle r_{ii} \\
& \quad + (y_{ti} - \langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_i)\sum_{j=1,j\neq i}^{n}(y_{tj} - \langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_j)\left\langle\frac{1}{\sqrt{B_i\sigma_i B_j\sigma_j}}\right\rangle\left\langle\frac{1}{\sqrt{v_{ti}v_{tj}}}\right\rangle r_{ij}\Bigg) \\
& + \sum_{i=1}^{n}\Bigg(\left\langle\frac{C_i|\gamma_i|}{B_i}\right\rangle\left\langle\frac{1}{v_{ti}}\right\rangle\langle s_{ti}\rangle(y_{ti} - \langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_i)r_{ii} \\
& \quad + \langle s_{ti}\rangle\sum_{j=1,j\neq i}^{n}\left\langle\frac{C_i|\gamma_i|\sqrt{\sigma_i}}{\sqrt{B_iB_j\sigma_j}}\right\rangle(y_{tj} - \langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_j)\left\langle\frac{1}{\sqrt{v_{ti}v_{tj}}}\right\rangle r_{ij}\Bigg) \\
& - \frac{1}{2}\sum_{i=1}^{n}\Bigg(\left\langle\frac{C_i^2|\gamma_i|^2\sigma_i}{B_i}\right\rangle\langle s_{ti}^2\rangle\left\langle\frac{1}{v_{ti}}\right\rangle r_{ii} \\
& \quad + \sum_{j=1,j\neq i}^{n}\left\langle\frac{C_iC_j|\gamma_i\gamma_j|\sqrt{\sigma_i\sigma_j}}{\sqrt{B_iB_j}}\right\rangle\langle s_{ti}s_{tj}\rangle\left\langle\frac{1}{\sqrt{v_{ti}v_{tj}}}\right\rangle r_{ij}\Bigg) \\
& + \sum_{i=1}^{n}\Bigg(\left\langle\frac{A_i}{B_i\sigma_i}\right\rangle(y_{ti} - \langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_i)r_{ii} \\
& \quad + \sum_{j=1,j\neq i}^{n}\left\langle\frac{A_i}{\sqrt{B_i\sigma_iB_j\sigma_j}}\right\rangle\left\langle\frac{\sqrt{v_{ti}}}{\sqrt{v_{tj}}}\right\rangle(y_{tj} - \langle \mathbf{F}_t'\boldsymbol{\theta}_t\rangle_j)r_{ij}\Bigg) \\
& - \sum_{i=1}^{n}\Bigg(\left\langle\frac{A_iC_i|\gamma_i|}{B_i}\right\rangle\langle s_{ti}\rangle r_{ii} \\
& \quad + \sum_{j=1,j\neq i}^{n}\left\langle\frac{A_iC_j|\gamma_j|\sqrt{\sigma_j}}{\sqrt{B_iB_j\sigma_i}}\right\rangle\langle s_{tj}\rangle\left\langle\frac{\sqrt{v_{ti}}}{\sqrt{v_{tj}}}\right\rangle r_{ij}\Bigg) \\
& - \frac{1}{2}\sum_{i=1}^{n}\Bigg(\left\langle\frac{A_i^2}{B_i\sigma_i}\right\rangle\langle v_{ti}\rangle r_{ii} \\
& \quad + \sum_{j=1,j\neq i}^{n}\left\langle\frac{A_iA_j}{\sqrt{B_i\sigma_iB_j\sigma_j}}\right\rangle\langle\sqrt{v_{ti}v_{tj}}\rangle r_{ij}\Bigg)\Bigg\}.
\end{aligned}
$$

Here $f_0(\boldsymbol{\sigma})$, $f_0(\boldsymbol{\gamma})$, and $f_0(\mathbf{R})$ denote the prior distributions of $\boldsymbol{\sigma}$, $\boldsymbol{\gamma}$, and $\mathbf{R}$, respectively. Thus, we can update any of these distributions with importance sampling as detailed in Section A.3

A.6 MVexDQLM ISVB Closed Form Integrals

For notational simplicity, we omit the superscript indicating the VB iteration.

If $r(\mathbf{s}_t) = \text{MVN}^+(\boldsymbol{\mu}_{s_t}, \boldsymbol{\Sigma}_{s_t})$, the expectations $\langle s_{ti} \rangle$, $\langle s_{ti}^2 \rangle$, $\langle s_{ti} s_{tj} \rangle$ for $i, j = 1, \ldots, n$ with $i \neq j$ are available in closed form as detailed in Manjunath and Wilhelm (2012) and can be calculated using the R package **tmvtnorm** (Wilhelm and G, 2015).

If $r(\boldsymbol{\theta}_t) = \text{N}(\mathbf{m}_t^s, \mathbf{C}_t^s)$, then

$$\begin{aligned} \langle \mathbf{F}_t' \boldsymbol{\theta}_t \rangle &= \mathbf{F}_t' \mathbf{m}_t^s \\ \langle (\mathbf{F}_t' \boldsymbol{\theta}_t)^2 \rangle &= \mathbf{F}_t' \mathbf{C}_t^s \mathbf{F}_t + (\mathbf{F}_t' \mathbf{m}_t^s)^2 \end{aligned}$$

If $r(\mathbf{v}_t)$, for $t = 1, \ldots, T$, is approximated with IS according to Equation (A.2), the following expectations for $i, j = 1, \ldots, n$ and $i \neq j$ can be approximated using Equation (A.3);

$$\langle \sqrt{v_{ti} v_{tj}} \rangle, \langle v_{ti} \rangle, \left\langle \sqrt{\frac{v_{ti}}{v_{tj}}} \right\rangle, \left\langle \frac{1}{\sqrt{v_{ti} v_{tj}}} \right\rangle, \left\langle \frac{1}{v_{ti}} \right\rangle.$$

Similarly, if $r(\boldsymbol{\sigma}, \boldsymbol{\gamma})$ is approximated with IS, the following expectations can be approximated using Equation (A.3);

$$\left\langle \frac{1}{B_i \sigma_i} \right\rangle, \left\langle \frac{1}{\sqrt{B_i \sigma_i B_j \sigma_j}} \right\rangle, \left\langle \frac{C_i |\gamma_i|}{B_i} \right\rangle, \left\langle \frac{C_i |\gamma_i| \sqrt{\sigma_i}}{\sqrt{B_i B_j \sigma_j}} \right\rangle, \left\langle \frac{C_i^2 |\gamma_i|^2 \sigma_i}{B_i} \right\rangle,$$

$$\left\langle \frac{C_i C_j |\gamma_i \gamma_j| \sqrt{\sigma_i \sigma_j}}{\sqrt{B_i B_j}} \right\rangle, \left\langle \frac{A_i}{B_i \sigma_i} \right\rangle, \left\langle \frac{A_i}{\sqrt{B_i \sigma_i B_j \sigma_j}} \right\rangle, \left\langle \frac{A_i C_i |\gamma_i|}{B_i} \right\rangle,$$

$$\left\langle \frac{A_i C_j |\gamma_j| \sqrt{\sigma_j}}{\sqrt{B_i B_j \sigma_i}} \right\rangle, \left\langle \frac{A_i^2}{B_i \sigma_i} \right\rangle, \left\langle \frac{A_i A_j}{\sqrt{B_i \sigma_i B_j \sigma_j}} \right\rangle.$$

Lastly, if $r(\mathbf{R})$ is approximated with IS, $\langle r_{ij} \rangle$ for $i, j = 1, \ldots, n$ can be approx-

imated using Equation (A.3) as well.

A.7 Multivariate IVT analysis additional results

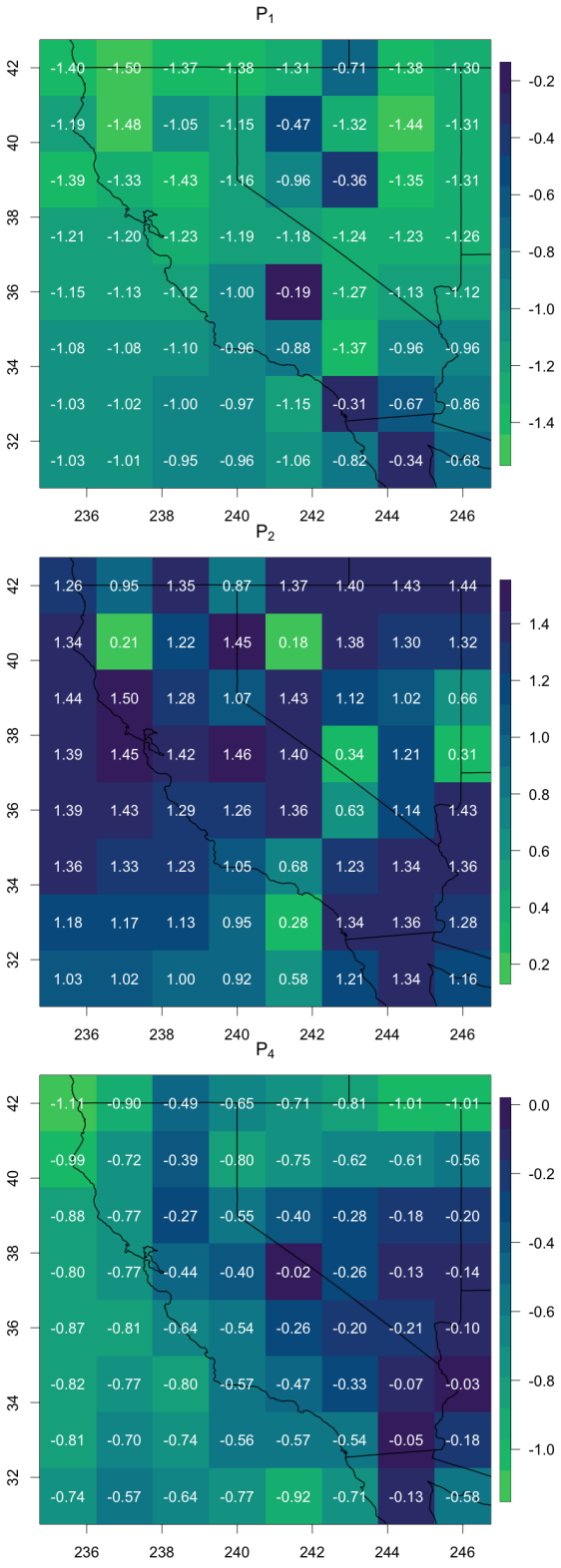

Figure A.1: MAP estimates of the annual (top), semi-annual (middle), and quarterly (bottom) phases displayed by latitude (y-axes) and longitude (x-axes).

	235.5	237	238.5	240	241.5	243	244.5	246
42	169.9 (168.6,171.3)	131.9 (130.8,133.0)	121.7 (120.8,122.7)	129.9 (128.9,130.9)	132.1 (131.1,133.1)	131.2 (130.3,132.1)	137.0 (136.2,137.9)	135.8 (135.0,136.7)
40.5	169.5 (168.3,170.8)	128.6 (127.5,129.6)	119.1 (118.2,120.1)	120.8 (119.9,121.8)	115.7 (114.8,116.6)	114.2 (113.4,115.0)	120.9 (120.1,121.7)	122.7 (121.9,123.5)
39	176.3 (175.1,177.6)	138.6 (137.5,139.7)	117.0 (116.0,118.0)	111.1 (110.2,112.0)	99.9 (99.2,100.6)	107.7 (107.0,108.4)	117.9 (117.2,118.7)	121.9 (121.1,122.7)
37.5	181.1 (179.9,182.4)	150.7 (149.5,151.8)	121.9 (120.9,122.8)	92.9 (92.2,93.7)	85.0 (84.4,85.6)	93.9 (93.3,94.6)	114.6 (113.8,115.3)	119.0 (118.2,119.8)
36	174.7 (173.5,175.9)	155.6 (154.5,156.7)	130.7 (129.7,131.7)	104.2 (103.4,105.0)	93.7 (93.0,94.4)	100.1 (99.4,100.7)	115.1 (114.3,115.8)	113.5 (112.7,114.3)
34.5	169.0 (167.9,170.1)	156.7 (155.6,157.8)	139.4 (138.5,140.4)	122.9 (122.0,123.8)	108.1 (107.3,108.9)	111.7 (110.9,112.4)	114.4 (113.6,115.2)	113.1 (112.3,113.9)
33	163.4 (162.3,164.5)	157.0 (155.9,158.0)	149.3 (148.3,150.3)	134.9 (134.0,135.9)	116.1 (115.2,117.0)	107.7 (106.9,108.5)	106.4 (105.6,107.2)	116.2 (115.4,117.0)
31.5	153.9 (152.8,154.9)	148.1 (147.2,149.1)	143.6 (142.6,144.6)	138.5 (137.6,139.5)	126.8 (125.9,127.7)	110.6 (109.7,111.4)	105.6 (104.8,106.4)	113.1 (112.3,113.9)

Table A.1: MAP and 95% CrI estimates of the second-order polynomial trend component, $\eta_{1,1}$, indexed by latitude (rows) and longitude (columns).

	235.5	237	238.5	240	241.5	243	244.5	246
42	22.84 (20.46,25.22)	11.52 (9.60,13.44)	9.56 (7.87,11.26)	7.44 (5.72,9.17)	3.91 (2.23,5.60)	0.40 (-1.16,1.96)	-1.88 (-3.37,-0.38)	-2.05 (-3.52,-0.58)
40.5	15.62 (13.41,17.82)	8.40 (6.57,10.24)	9.27 (7.60,10.95)	6.53 (4.87,8.19)	2.13 (0.61,3.66)	-1.06 (-2.45,0.33)	-2.66 (-4.00,-1.32)	-1.82 (-3.17,-0.46)
39	13.82 (11.64,16.01)	6.40 (4.47,8.32)	8.17 (6.46,9.88)	4.91 (3.34,6.47)	0.96 (-0.29,2.21)	-0.16 (-1.38,1.06)	0.12 (-1.18,1.42)	2.91 (1.53,4.29)
37.5	18.01 (15.82,20.20)	8.58 (6.61,10.56)	4.42 (2.71,6.13)	1.87 (0.56,3.18)	1.65 (0.58,2.72)	2.38 (1.31,3.46)	3.80 (2.50,5.10)	3.75 (2.35,5.14)
36	21.04 (18.96,23.11)	11.01 (9.10,12.92)	4.27 (2.57,5.96)	0.22 (-1.15,1.60)	-1.30 (-2.48,-0.13)	1.36 (0.17,2.54)	5.15 (3.79,6.50)	6.82 (5.41,8.23)
34.5	19.95 (17.99,21.91)	12.87 (11.03,14.70)	5.23 (3.52,6.95)	-2.83 (-4.37,-1.28)	-5.14 (-6.54,-3.74)	0.06 (-1.32,1.43)	6.24 (4.89,7.60)	7.68 (6.30,9.07)
33	16.78 (14.89,18.66)	14.50 (12.71,16.29)	9.28 (7.56,11.00)	-1.86 (-3.48,-0.23)	-7.53 (-9.03,-6.03)	-0.54 (-1.94,0.86)	6.60 (5.27,7.94)	6.85 (5.44,8.25)
31.5	13.80 (12.01,15.60)	13.88 (12.16,15.59)	13.90 (12.23,15.57)	8.54 (6.90,10.19)	0.84 (-0.76,2.45)	1.30 (-0.17,2.77)	5.85 (4.45,7.25)	6.61 (5.18,8.04)

Table A.2: MAP and 95% CrI estimates of the second-order polynomial trend component, η_2, indexed by latitude (rows) and longitude (columns).

	235.5	237	238.5	240	241.5	243	244.5	246
42	24.36 (22.97,25.75)	20.36 (19.25,21.47)	9.23 (8.26,10.21)	5.42 (4.42,6.41)	2.93 (1.97,3.90)	1.35 (0.45,2.24)	6.29 (5.44,7.15)	11.91 (11.07,12.75)
40.5	19.57 (18.29,20.85)	20.67 (19.61,21.73)	7.83 (6.86,8.79)	2.73 (1.78,3.68)	1.23 (0.35,2.10)	3.48 (2.69,4.28)	6.43 (5.67,7.20)	12.18 (11.41,12.96)
39	17.39 (16.13,18.66)	22.66 (21.55,23.77)	12.47 (11.48,13.45)	4.86 (3.97,5.76)	1.70 (0.99,2.41)	5.18 (4.49,5.87)	10.83 (10.09,11.57)	17.13 (16.34,17.92)
37.5	14.31 (13.04,15.57)	20.84 (19.70,21.98)	17.41 (16.42,18.39)	7.43 (6.68,8.18)	2.38 (1.78,2.99)	7.08 (6.48,7.69)	15.57 (14.82,16.31)	17.85 (17.05,18.65)
36	14.90 (13.70,16.10)	19.25 (18.15,20.36)	18.12 (17.14,19.09)	9.32 (8.53,10.10)	2.63 (1.96,3.30)	6.16 (5.48,6.83)	14.76 (13.99,15.53)	17.38 (16.57,18.19)
34.5	16.01 (14.88,17.14)	18.22 (17.16,19.28)	18.63 (17.65,19.62)	13.40 (12.51,14.29)	4.40 (3.60,5.20)	4.61 (3.82,5.39)	9.25 (8.47,10.03)	14.75 (13.96,15.54)
33	15.75 (14.66,16.84)	16.42 (15.39,17.45)	16.95 (15.96,17.95)	15.40 (14.47,16.34)	8.19 (7.33,9.05)	0.85 (0.05,1.65)	7.41 (6.65,8.17)	16.14 (15.34,16.95)
31.5	13.87 (12.84,14.91)	14.85 (13.86,15.84)	14.09 (13.13,15.05)	14.29 (13.34,15.24)	11.07 (10.16,11.99)	3.42 (2.58,4.26)	6.46 (5.66,7.26)	18.11 (17.29,18.93)

Table A.3: MAP and 95% CrI estimates of the annual amplitude, A_1^j, indexed by latitude (rows) and longitude (columns).

	235.5	237	238.5	240	241.5	243	244.5	246
42	11.51 (10.13,12.89)	10.83 (9.73,11.93)	7.95 (6.98,8.91)	8.40 (7.41,9.38)	8.34 (7.38,9.30)	7.75 (6.87,8.64)	7.49 (6.64,8.34)	5.10 (4.27,5.93)
40.5	13.87 (12.60,15.15)	11.61 (10.56,12.66)	6.49 (5.53,7.45)	7.14 (6.20,8.09)	6.81 (5.95,7.68)	6.21 (5.43,7.00)	5.57 (4.81,6.32)	2.61 (1.84,3.37)
39	15.32 (14.06,16.57)	13.57 (12.47,14.67)	5.56 (4.59,6.54)	5.76 (4.87,6.65)	5.64 (4.94,6.34)	4.57 (3.89,5.25)	1.58 (0.85,2.31)	1.14 (0.36,1.92)
37.5	14.39 (13.13,15.65)	13.70 (12.57,14.84)	8.58 (7.61,9.55)	5.07 (4.33,5.82)	4.31 (3.72,4.91)	0.32 (-0.27,0.92)	3.91 (3.18,4.64)	3.73 (2.94,4.52)
36	13.40 (12.21,14.59)	12.70 (11.60,13.79)	10.84 (9.87,11.81)	7.17 (6.39,7.95)	3.65 (2.99,4.31)	1.92 (1.26,2.59)	6.33 (5.56,7.09)	6.38 (5.58,7.18)
34.5	12.37 (11.25,13.49)	11.65 (10.60,12.70)	10.01 (9.03,10.99)	7.37 (6.49,8.25)	1.73 (0.94,2.52)	2.52 (1.74,3.29)	6.68 (5.91,7.45)	9.11 (8.33,9.90)
33	10.83 (9.75,11.91)	10.94 (9.92,11.97)	9.75 (8.77,10.73)	6.92 (5.99,7.84)	2.01 (1.16,2.86)	3.36 (2.56,4.15)	7.71 (6.95,8.46)	11.74 (10.94,12.54)
31.5	8.60 (7.58,9.63)	8.77 (7.79,9.75)	8.60 (7.65,9.55)	7.49 (6.55,8.43)	3.56 (2.65,4.47)	3.11 (2.28,3.95)	9.08 (8.29,9.88)	14.58 (13.77,15.39)

Table A.4: MAP and 95% CrI estimates of the semi-annual amplitude, A_2^j, indexed by latitude (rows) and longitude (columns).

	235.5	237	238.5	240	241.5	243	244.5	246
42	4.97 (3.59,6.34)	3.63 (2.53,4.73)	2.78 (1.82,3.74)	3.70 (2.72,4.67)	3.97 (3.02,4.93)	3.47 (2.59,4.36)	3.27 (2.42,4.12)	2.89 (2.07,3.72)
40.5	6.13 (4.86,7.39)	3.88 (2.83,4.92)	3.00 (2.05,3.96)	3.94 (3.00,4.88)	3.68 (2.82,4.54)	2.84 (2.06,3.63)	2.52 (1.76,3.27)	2.71 (1.95,3.47)
39	6.96 (5.71,8.21)	4.54 (3.44,5.63)	3.04 (2.07,4.01)	3.24 (2.36,4.13)	2.74 (2.05,3.44)	2.52 (1.84,3.20)	2.52 (1.79,3.24)	2.94 (2.16,3.71)
37.5	7.37 (6.11,8.62)	5.19 (4.06,6.32)	3.66 (2.69,4.62)	2.91 (2.18,3.64)	2.31 (1.72,2.90)	2.35 (1.76,2.94)	2.77 (2.04,3.49)	3.35 (2.57,4.14)
36	6.85 (5.67,8.04)	5.35 (4.26,6.44)	4.10 (3.14,5.06)	3.15 (2.38,3.92)	2.50 (1.85,3.16)	2.54 (1.88,3.19)	2.98 (2.22,3.74)	3.51 (2.71,4.31)
34.5	5.93 (4.81,7.05)	4.89 (3.84,5.94)	3.59 (2.62,4.56)	2.86 (1.99,3.73)	2.52 (1.73,3.30)	2.60 (1.83,3.37)	2.64 (1.88,3.41)	2.46 (1.67,3.24)
33	4.70 (3.62,5.77)	4.19 (3.17,5.21)	3.55 (2.57,4.53)	2.73 (1.81,3.65)	2.36 (1.51,3.21)	1.62 (0.84,2.41)	1.38 (0.63,2.13)	1.29 (0.49,2.08)
31.5	3.65 (2.63,4.68)	3.14 (2.16,4.12)	3.09 (2.14,4.04)	2.97 (2.03,3.90)	2.83 (1.92,3.73)	2.22 (1.39,3.05)	1.46 (0.67,2.25)	1.40 (0.60,2.21)

Table A.5: MAP and 95% CrI estimates of the quarterly amplitude, A_4^j, indexed by latitude (rows) and longitude (columns).

	235.5	237	238.5	240	241.5	243	244.5	246
42	-1.40 (-1.44,-1.36)	-1.50 (-2.10,-0.89)	-1.37 (-1.45,-1.30)	-1.38 (-1.51,-1.24)	-1.31 (-1.90,-0.71)	-0.71 (-1.24,-0.18)	-1.38 (-1.48,-1.27)	-1.30 (-1.34,-1.25)
40.5	-1.19 (-3.10,0.71)	-1.48 (-1.52,-1.45)	-1.05 (-3.21,1.10)	-1.15 (-1.41,-0.90)	-0.47 (-0.94,0.01)	-1.32 (-2.32,-0.32)	-1.44 (-2.04,-0.84)	-1.31 (-1.36,-1.27)
39	-1.39 (-1.44,-1.34)	-1.33 (-1.37,-1.30)	-1.43 (-1.49,-1.38)	-1.16 (-1.27,-1.05)	-0.96 (-1.25,-0.67)	-0.36 (-3.30,2.57)	-1.35 (-1.40,-1.30)	-1.31 (-1.34,-1.27)
37.5	-1.21 (-1.27,-1.15)	-1.20 (-1.23,-1.16)	-1.23 (-1.27,-1.19)	-1.19 (-1.26,-1.11)	-1.18 (-1.37,-0.99)	-1.24 (-1.30,-1.19)	-1.23 (-1.27,-1.20)	-1.26 (-1.29,-1.22)
36	-1.15 (-1.20,-1.10)	-1.13 (-1.17,-1.09)	-1.12 (-1.15,-1.08)	-1.00 (-1.05,-0.95)	-0.19 (-0.38,-0.00)	-1.27 (-1.35,-1.20)	-1.13 (-1.17,-1.10)	-1.12 (-1.16,-1.08)
34.5	-1.08 (-1.12,-1.03)	-1.08 (-1.12,-1.04)	-1.10 (-1.13,-1.06)	-0.96 (-1.01,-0.91)	-0.88 (-1.01,-0.74)	-1.37 (-1.49,-1.26)	-0.96 (-1.02,-0.91)	-0.96 (-0.99,-0.92)
33	-1.03 (-1.08,-0.98)	-1.02 (-1.07,-0.98)	-1.00 (-1.04,-0.95)	-0.97 (-1.02,-0.93)	-1.15 (-1.21,-1.09)	-0.31 (-1.20,0.57)	-0.67 (-0.75,-0.59)	-0.86 (-0.89,-0.82)
31.5	-1.03 (-1.08,-0.98)	-1.01 (-1.05,-0.97)	-0.95 (-1.00,-0.91)	-0.96 (-1.01,-0.91)	-1.06 (-1.12,-1.01)	-0.82 (-3.26,1.63)	-0.34 (-0.43,-0.26)	-0.68 (-0.72,-0.64)

Table A.6: MAP and 95% CrI estimates of the annual phase, P_1^j, indexed by latitude (rows) and longitude (columns).

	235.5	237	238.5	240	241.5	243	244.5	246
42	1.26 (1.17,1.35)	0.95 (-1.44,3.33)	1.35 (1.27,1.43)	0.87 (-1.59,3.33)	1.37 (0.19,2.55)	1.40 (0.56,2.23)	1.43 (0.84,2.02)	1.44 (0.84,2.04)
40.5	1.34 (1.27,1.40)	0.21 (-2.80,3.23)	1.22 (1.11,1.32)	1.45 (1.35,1.54)	0.18 (-2.81,3.18)	1.38 (1.31,1.45)	1.30 (1.20,1.40)	1.32 (1.13,1.51)
39	1.44 (1.38,1.51)	1.50 (1.45,1.55)	1.28 (1.14,1.42)	1.07 (-1.00,3.15)	1.43 (1.35,1.52)	1.12 (1.01,1.23)	1.02 (0.69,1.36)	0.66 (-1.65,2.97)
37.5	1.39 (0.20,2.58)	1.45 (1.40,1.51)	1.42 (1.34,1.50)	1.46 (1.37,1.56)	1.40 (1.31,1.50)	0.34 (-1.22,1.90)	1.21 (1.09,1.33)	0.31 (-2.59,3.22)
36	1.39 (0.20,2.58)	1.43 (1.37,1.50)	1.29 (1.22,1.35)	1.26 (1.17,1.34)	1.36 (1.24,1.47)	0.63 (0.39,0.87)	1.14 (1.04,1.23)	1.43 (1.33,1.52)
34.5	1.36 (1.30,1.43)	1.33 (1.27,1.39)	1.23 (1.16,1.30)	1.05 (0.97,1.13)	0.68 (0.35,1.01)	1.23 (-0.14,2.60)	1.34 (1.26,1.42)	1.36 (1.29,1.43)
33	1.18 (1.11,1.26)	1.17 (1.10,1.23)	1.13 (1.06,1.20)	0.95 (0.85,1.06)	0.28 (-0.05,0.62)	1.34 (1.15,1.54)	1.36 (1.29,1.42)	1.28 (1.23,1.33)
31.5	1.03 (0.93,1.13)	1.02 (0.94,1.11)	1.00 (0.93,1.08)	0.92 (0.84,1.00)	0.58 (0.40,0.76)	1.21 (1.01,1.41)	1.34 (1.28,1.39)	1.16 (1.12,1.20)

Table A.7: MAP and 95% CrI estimates of the semi-annual phase, P_2^j, indexed by latitude (rows) and longitude (columns).

	235.5	237	238.5	240	241.5	243	244.5	246
42	-1.11 (-1.32,-0.89)	-0.90 (-1.12,-0.69)	-0.49 (-0.74,-0.25)	-0.65 (-0.83,-0.46)	-0.71 (-0.89,-0.53)	-0.81 (-1.00,-0.61)	-1.01 (-1.20,-0.83)	-1.01 (-1.21,-0.82)
40.5	-0.99 (-1.13,-0.85)	-0.72 (-0.92,-0.52)	-0.39 (-0.62,-0.16)	-0.80 (-0.95,-0.66)	-0.75 (-0.90,-0.61)	-0.62 (-0.81,-0.43)	-0.61 (-0.83,-0.39)	-0.56 (-0.77,-0.35)
39	-0.88 (-1.00,-0.75)	-0.77 (-0.96,-0.58)	-0.27 (-0.49,-0.06)	-0.55 (-0.75,-0.35)	-0.40 (-0.60,-0.20)	-0.28 (-0.48,-0.08)	-0.18 (-0.39,0.02)	-0.20 (-0.39,-0.00)
37.5	-0.80 (-0.93,-0.67)	-0.77 (-0.92,-0.62)	-0.44 (-0.64,-0.23)	-0.40 (-0.57,-0.23)	-0.02 (-0.24,0.19)	-0.26 (-0.42,-0.10)	-0.13 (-0.29,0.02)	-0.14 (-0.30,0.02)
36	-0.87 (-0.99,-0.75)	-0.81 (-0.98,-0.65)	-0.64 (-0.80,-0.49)	-0.54 (-0.72,-0.35)	-0.26 (-0.45,-0.07)	-0.20 (-0.40,-0.01)	-0.21 (-0.39,-0.04)	-0.10 (-0.26,0.07)
34.5	-0.82 (-0.94,-0.69)	-0.77 (-0.92,-0.62)	-0.80 (-0.97,-0.63)	-0.57 (-0.84,-0.31)	-0.47 (-0.68,-0.27)	-0.33 (-0.52,-0.14)	-0.07 (-0.26,0.12)	-0.03 (-0.28,0.21)
33	-0.81 (-0.98,-0.65)	-0.70 (-0.88,-0.52)	-0.74 (-0.94,-0.53)	-0.56 (-0.77,-0.34)	-0.57 (-0.83,-0.31)	-0.54 (-0.85,-0.23)	-0.05 (-0.37,0.26)	-0.18 (-0.63,0.27)
31.5	-0.74 (-0.96,-0.52)	-0.57 (-0.76,-0.38)	-0.64 (-0.86,-0.42)	-0.77 (-1.01,-0.54)	-0.92 (-1.15,-0.69)	-0.71 (-0.97,-0.45)	-0.13 (-0.49,0.23)	-0.58 (-0.99,-0.17)

Table A.8: MAP and 95% CrI estimates of the quarterly phase, P_4^j, indexed by latitude (rows) and longitude (columns).

Bibliography

Tracy M Backes, Michael L Kaplan, Rina Schumer, and John F Mejia. A climatology of the vertical structure of water vapor transport to the Sierra Nevada in cool season atmospheric river precipitation events. *Journal of Hydrometeorology*, 16(3):1029–1047, 2015.

Raquel Barata, Raquel Prado, and Bruno Sansó. Supplement to "Fast inference for time-varying quantiles via flexible dynamic models with application to the characterization of atmospheric rivers". *The Annals of Applied Statistics*, 2021.

Raquel Barata, Raquel Prado, and Bruno Sansó. Fast inference for time-varying quantiles via flexible dynamic models with application to the characterization of atmospheric rivers. *The Annals of Applied Statistics*, in press.

David Barber and Silvia Chiappa. Unified inference for variational Bayesian linear Gaussian state-space models. In *Advances in Neural Information Processing Systems*, pages 81–88, 2007.

Matthew J Beal. *Variational algorithms for approximate Bayesian inference*. PhD thesis, UCL (University College London), 2003.

Dries Benoit, Rahim Al-Hamzawi, Keming Yu, and Dirk Van den Poel. ***bayesQR**: A Bayesian approach to quantile regression*, 2011. URL https://CRAN.R-project.org/package=bayesQR. R package version 2.3.

Dries F Benoit and Dirk Van den Poel. bayesqr: A Bayesian approach to quantile regression. *Journal of Statistical Software*, 76(1):1–32, 2017.

Mauro Bernardi, Roberto Casarin, Bertrand Maillet, and Lea Petrella. Dynamic model averaging for Bayesian quantile regression. *arXiv: Statistics Theory*, 2016.

Paul Berrisford, P Kållberg, S Kobayashi, D Dee, S Uppala, AJ Simmons, P Poli, and H Sato. Atmospheric conservation properties in ERA-interim. *Quarterly Journal of the Royal Meteorological Society*, 137(659):1381–1399, 2011.

Lindsay R Berry and Mike West. Bayesian forecasting of many count-valued time series. *Journal of Business & Economic Statistics*, 38(4):872–887, 2020.

David M Blei and John D Lafferty. Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning*, pages 113–120. ACM, 2006.

David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112 (518):859–877, 2017.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

Chris K Carter and Robert Kohn. On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553, 1994.

Wilson Ye Chen, Gareth W Peters, Richard H Gerlach, and Scott A Sisson. Dynamic quantile function models. *arXiv preprint arXiv:1707.02587*, 2017.

Dick P Dee, SM Uppala, AJ Simmons, Paul Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo, P Bauer, et al. The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597, 2011.

Michael Dettinger. Climate change, atmospheric rivers, and floods in California–a multimodel analysis of storm frequency and magnitude changes. *JAWRA Journal of the American Water Resources Association*, 47(3):514–523, 2011.

Michael D Dettinger. Atmospheric rivers as drought busters on the US West Coast. *Journal of Hydrometeorology*, 14(6):1721–1732, 2013.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

Reeves Fletcher and Colin M Reeves. Function minimization by conjugate gradients. *The Computer Journal*, 7(2):149–154, 1964.

Nick Foti, Jason Xu, Dillon Laird, and Emily Fox. Stochastic variational inference for hidden markov models. In *Advances in Neural Information Processing Systems*, pages 3599–3607, 2014.

Karen Fox. Solar minimum; solar maximum, 2012. URL https://www.nasa.gov/mission_pages/sunearth/news/solarmin-max.html.

Sylvia Frühwirth-Schnatter. Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202, 1994.

Dani Gamerman and Helio S Migon. Dynamic hierarchical models. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 629–642, 1993.

Dani Gamerman, Ajax RB Moreira, and Håvard Rue. Space-varying regression models: specifications and simulation. *Computational Statistics & Data Analysis*, 42(3):513–533, 2003.

Alan E Gelfand and Sujit K Ghosh. Model choice: a minimum posterior predictive loss approach. *Biometrika*, 85(1):1–11, 1998.

Kelly CM Gonçalves, Helio S Migon, and Leonardo S Bastos. Dynamic quantile linear models: A bayesian approach. *Bayesian Analysis*, 15(2):335–362, 2020.

Bin Guan and Duane E Waliser. Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *Journal of Geophysical Research: Atmospheres*, 120(24):12514–12535, 2015.

Bin Guan, Noah P Molotch, Duane E Waliser, Eric J Fetzer, and Paul J Neiman. Extreme snowfall events linked to atmospheric rivers and surface air temperature via satellite measurements. *Geophysical Research Letters*, 37(20), 2010.

Bin Guan, Duane E Waliser, Noah P Molotch, Eric J Fetzer, and Paul J Neiman. Does the Madden–Julian oscillation influence wintertime atmospheric rivers and snowpack in the Sierra Nevada? *Monthly Weather Review*, 140(2):325–342, 2012.

Bin Guan, Noah P Molotch, Duane E Waliser, Eric J Fetzer, and Paul J Neiman. The 2010/2011 snow season in California's Sierra Nevada: Role of atmospheric rivers and modes of large-scale variability. *Water Resources Research*, 49(10): 6731–6743, 2013.

Timothy Hanson and Wesley O Johnson. Modeling regression error with a mixture of Polya trees. *Journal of the American Statistical Association*, 97(460):1020–1033, 2002.

Norbert Henze. A probabilistic representation of the 'skew-normal' distribution. *Scandinavian Journal of Statistics*, pages 271–275, 1986.

Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

Darjus Hosszejni and Gregor Kastner. Approaches toward the Bayesian estimation of the stochastic volatility model with leverage. In *International Conference on Bayesian Statistics in Action*, pages 75–83. Springer, 2018.

Gabriel Huerta, Wenxin Jiang, and Martin A Tanner. Time series modeling via hierarchical mixtures. *Statistica Sinica*, pages 1097–1118, 2003.

Mary Johannis, Lorraine E Flint, Michael D Dettinger, Alan L Flint, and Regina Ochoa. The role of snowpack, rainfall, and reservoirs in buffering California against drought effects. Technical report, US Geological Survey, 2016.

Matthew Johnson and Alan Willsky. Stochastic variational inference for Bayesian time series models. In *International Conference on Machine Learning*, pages 1854–1862. PMLR, 2014.

Gregor Kastner. Dealing with stochastic volatility in time series using the R package stochvol. *Journal of Statistical Software, Articles*, 69(5):1–30, 2016. ISSN 1548-7660. doi: 10.18637/jss.v069.i05. URL https://www.jstatsoft.org/v069/i05.

Matthias Katzfuss, Jonathan R Stroud, and Christopher K Wikle. Understanding the ensemble Kalman filter. *The American Statistician*, 70(4):350–357, 2016.

Matthias Katzfuss, Jonathan R Stroud, and Christopher K Wikle. Ensemble kalman methods for high-dimensional hierarchical dynamic space-time models. *Journal of the American Statistical Association*, 115(530):866–885, 2020.

Roger Koenker. *Quantile regression*. Cambridge University Press, New York, 2005.

Roger Koenker. ***quantreg**: Quantile Regression*, 2021. URL https://CRAN.R-project.org/package=quantreg. R package version 5.85.

Roger Koenker and Zhijie Xiao. Quantile autoregression. *Journal of the American Statistical Association*, 101(475):980–990, 2006.

Ivana Komunjer. Quasi-maximum likelihood estimation for conditional quantiles. *Journal of Econometrics*, 128(1):137–164, 2005.

Athanasios Kottas and Alan E Gelfand. Bayesian semiparametric median regression modeling. *Journal of the American Statistical Association*, 96(456): 1458–1468, 2001.

Athanasios Kottas and Milovan Krnjajić. Bayesian semiparametric modelling in quantile regression. *Scandinavian Journal of Statistics*, 36(2):297–319, 2009.

Samuel Kotz, Tomasz Kozubowski, and Krzystof Podgorski. *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance*. Springer Science & Business Media, 2001.

Samuel Kotz, Tomaz J Kozubowski, and Krzysztof Podgórski. Asymmetric multivariate Laplace distribution. In *The Laplace Distribution and Generalizations*, pages 239–272. Springer, 2003.

Hideo Kozumi and Genya Kobayashi. Gibbs sampling methods for Bayesian quantile regression. *Journal of Statistical Computation and Simulation*, 81(11):1565–1578, 2011.

Solomon Kullback and Richard A Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

Brunero Liseo and Nicola Loperfido. A note on reference priors for the scalar skew-normal distribution. *Journal of Statistical Planning and Inference*, 136 (2):373–389, 2006.

Pamela J Lombard. Flood-and drought-related natural hazards activities of the US Geological Survey in New England. Technical report, US Geological Survey, 2016.

Kristian Lum, Alan E Gelfand, et al. Spatial quantile multiple regression using the asymmetric Laplace process. *Bayesian Analysis*, 7(2):235–258, 2012.

BG Manjunath and Stefan Wilhelm. Moments calculation for the doubly truncated multivariate normal density. *arXiv preprint arXiv:1206.5387*, 2012.

Helio S Migon and Larissa C Alves. Multivariate dynamic regression: Modeling and forecasting for intraday electricity load. *Applied Stochastic Models in Business and Industry*, 29(6):579–598, 2013.

Brian Neelon, Fan Li, Lane F Burgette, and Sara E Benjamin Neelon. A spatiotemporal quantile regression model for emergency department expenditures. *Statistics in Medicine*, 34(17):2559–2575, 2015.

Paul J Neiman, Allen B White, F Martin Ralph, Daniel J Gottas, and Seth I Gutman. A water vapour flux tool for precipitation forecasting. In *Proceedings of the Institution of Civil Engineers-Water Management*, volume 162, pages 83–94. Thomas Telford Ltd, 2009.

John A Nelder and Roger Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

NOAA National Centers for Environmental Information. Streamflow drought indicators, 2021. URL https://www.ncdc.noaa.gov/monitoring-references/dyk/streamflow-indicators.

Thomas Opitz, Raphaël Huser, Haakon Bakka, and Haavard Rue. INLA goes extreme: Bayesian tail regression for the estimation of high spatio-temporal quantiles. *Extremes*, pages 1–22, 2018.

Andrea Christina Ostroff, Clint C Muhlfeld, Patrick M Lambert, Nathaniel L Booth, Shawn L Carter, Jason M Stoker, and Michael J Focazio. USGS integrated drought science. *US Geological Survey Circular*, (1430), 2017.

Dirk Ostwald, Evgeniya Kirilina, Ludger Starke, and Felix Blankenburg. A tutorial on variational Bayes for latent linear stochastic time-series models. *Journal of Mathematical Psychology*, 60:1–19, 2014.

Lu Ou, Michael Hunter, Sy-Min Chow, Linying Ji, Meng Chen, Hui-Ju Hung, Jungmin Lee, Yanling Li, and Jonathan Park. ***dynr****: Dynamic Models with Regime-Switching*, 2021. URL https://CRAN.R-project.org/package=dynr. R package version 0.1.16-2.

Tullia Padellini and Haavard Rue. Model-aware quantile regression for discrete data. *arXiv preprint arXiv:1804.03714*, 2018.

Florentina Paraschiv, Derek W Bunn, and Sjur Westgaard. Estimation and application of fully parametric multifactor quantile regression with dynamic coefficients. *University of St. Gallen, School of Finance Research Paper*, 2016.

Christina M Patricola, John P O'Brien, Mark D Risser, Alan M Rhoades, Travis A O'Brien, Paul A Ullrich, Dáithí A Stone, and William D Collins. Maximizing ENSO as a source of western us hydroclimate predictability. *Climate Dynamics*, 54(1-2):351–372, 2020.

Will Penny, Stefan Kiebel, and Karl Friston. Variational Bayesian inference for fMRI time series. *NeuroImage*, 19(3):727–741, 2003.

Giovanni Petris and R An. An R package for dynamic linear models. *Journal of Statistical Software*, 36(12):1–16, 2010.

Giovanni Petris and Wally Gilks. ***dlm****: Bayesian and Likelihood Analysis of Dynamic Linear Models*, 2018. URL https://CRAN.R-project.org/package=dlm. R package version 1.1-5.

Raquel Prado and Mike West. *Time series: modeling, computation, and inference*. CRC Press, 2010.

Raquel Prado, Francisco Molina, and Gabriel Huerta. Multivariate time series modeling and classification via hierarchical VAR mixtures. *Computational Statistics & Data Analysis*, 51(3):1445–1462, 2006.

Jose M Quintana and Mike West. An analysis of international exchange rates using multivariate DLM's. *The Statistician*, pages 275–281, 1987.

Matias Quiroz, David J Nott, and Robert Kohn. Gaussian variational approximation for high-dimensional state space models. *arXiv preprint arXiv:1801.07873*, 2018.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL http://www.R-project.org/.

FM Ralph and MD Dettinger. Storms, floods, and the science of atmospheric rivers. *Eos, Transactions American Geophysical Union*, 92(32):265–266, 2011.

NAA Rayner, De E Parker, EB Horton, Chris K Folland, Lisa V Alexander, DP Rowell, Elizabeth C Kent, and A Kaplan. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, 108(D14), 2003.

William J Reed. The normal-Laplace distribution and its relatives. In *Advances in Distribution Theory, Order Statistics, and Inference*, pages 61–74. Springer, 2006.

Brian J Reich. Spatiotemporal quantile regression for detecting distributional changes in environmental processes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 61(4):535–553, 2012.

Brian J Reich and Luke B Smith. Bayesian quantile regression for censored data. *Biometrics*, 69(3):651–660, 2013.

Brian J Reich, Howard D Bondell, and Huixia J Wang. Flexible Bayesian quantile regression for independent and clustered data. *Biostatistics*, 11(2):337–352, 2009.

Brian J Reich, Montserrat Fuentes, and David B Dunson. Bayesian spatial quantile regression. *Journal of the American Statistical Association*, 106(493):6–20, 2011.

Murray Rosenblatt. Remarks on a multivariate transformation. *The Annals of Mathematical Statistics*, 23(3):470–472, 1952.

Jonathan J Rutz, W James Steenburgh, and F Martin Ralph. Climatological characteristics of atmospheric rivers and their inland penetration over the western United States. *Monthly Weather Review*, 142(2):905–921, 2014.

Simo Särkkä. *Bayesian filtering and smoothing*, volume 3. Cambridge University Press, 2013.

Matthew A Taddy and Athanasios Kottas. A Bayesian nonparametric approach to inference for quantile regression. *Journal of Business & Economic Statistics*, 28(3):357–369, 2010.

Surya T Tokdar, Joseph B Kadane, et al. Simultaneous linear quantile regression: A semiparametric Bayesian approach. *Bayesian Analysis*, 7(1):51–72, 2012.

Efthymios G Tsionas. Bayesian quantile inference. *Journal of Statistical Computation and Simulation*, 73(9):659–674, 2003.

Mark Tuckerman. *Statistical mechanics: Theory and molecular simulation*. Oxford University Press, 2010.

Rui Tuo. Uncertainty quantification with σ-stable process models. *Statistica Sinica*, 28:553–576, 2018.

Eli Tziperman, Mark A Cane, Stephen E Zebiak, Yan Xue, and B Blumenthal. Locking of El Nino's peak time to the end of the calendar year in the delayed oscillator picture of ENSO. *Journal of Climate*, 11(9):2191–2199, 1998.

U.S. Geological Survey. National water information system data available on the world wide web (usgs water data for the nation), 2016. URL http://waterdata.usgs.gov/nwis/.

U.S. National Integrated Drought Information System. National Integrated Drought Information System NIDIS - Drought.gov - U.S. Drought Portal, 2008. URL https://www.loc.gov/item/lcwaN0014288/.

Duane E Waliser, Mitchell W Moncrieff, David Burridge, Andreas H Fink, Dave Gochis, BN Goswami, Bin Guan, Patrick Harr, Julian Heming, Huang-Hsuing Hsu, et al. The year of tropical convection (May 2008–April 2010): Climate variability and weather highlights. *Bulletin of the American Meteorological Society*, 93(8):1189–1218, 2012.

Stephen Walker and Bani K Mallick. A Bayesian semiparametric accelerated failure time model. *Biometrics*, 55(2):477–483, 1999.

Grant B Weller, Daniel S Cooley, and Stephan R Sain. An investigation of the pineapple express phenomenon via bivariate extreme value theory. *Environmetrics*, 23(5):420–439, 2012.

Mike West and Jeff Harrison. *Bayesian forecasting and dynamic models*. Springer Science & Business Media, 2006.

Mike West, P Jeff Harrison, and Helio S Migon. Dynamic generalized linear models and Bayesian forecasting. *Journal of the American Statistical Association*, 80 (389):73–83, 1985.

Nuttanan Wichitaksorn, ST Boris Choy, and Richard Gerlach. A generalized class of skew distributions and associated robust quantile regression models. *Canadian Journal of Statistics*, 42(4):579–596, 2014.

Stefan Wilhelm and Manjunath B G. ***tmvtnorm**: Truncated Multivariate Normal and Student-t Distribution*, 2015. URL http://CRAN.R-project.org/package=tmvtnorm. R package version 1.4-10.

Ian N Williams and Christina M Patricola. Diversity of ENSO events unified by convective threshold sea surface temperature: A nonlinear ENSO index. *Geophysical Research Letters*, 45(17):9236–9244, 2018.

Yifei Yan. *Bayesian Modeling and Inference for Quantile Mixture Regression*. PhD thesis, University of California Santa Cruz, 2017.

Yifei Yan and Athanasios Kottas. A new family of error distributions for Bayesian quantile regression. *arXiv preprint arXiv:1701.05666*, 2017.

Keming Yu and Rana A Moyeed. Bayesian quantile regression. *Statistics & Probability Letters*, 54(4):437–447, 2001.

Dongming Zhu and John W Galbraith. Modeling and forecasting expected shortfall with the generalized asymmetric Student-t and asymmetric exponential power distributions. *Journal of Empirical Finance*, 18(4):765–778, 2011.

Dongming Zhu and Victoria Zinde-Walsh. Properties and estimation of asymmetric exponential power distribution. *Journal of Econometrics*, 148(1):86–99, 2009.